

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 6, 2002, 09:25:39 ; Search time 18.7 Seconds  
(without alignments)  
485,900 Million cell updates/sec

Title: US-09-892-287-1  
Perfect score: 372  
Sequence: 1 MDVLSPLFIKSHVHMOGI.....FOIPGSLREDPLGGAQPOI 372

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

W size: 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	372	100.0	372	2 US-08-884-681-1	Sequence 1, Appli
2	372	100.0	372	4 US-09-258-643-1	Sequence 1, Appli
3	295	79.3	329	2 US-08-884-681-3	Sequence 3, Appli
4	295	79.3	329	4 US-09-258-643-3	Sequence 3, Appli
5	8	2.2	946	3 US-08-560-005-4	Sequence 4, Appli
6	8	2.2	946	4 US-09-418-540-4	Sequence 4, Appli
7	8	1.9	206	6 5221624-30	Patent No. 5221624
8	7	1.9	207	2 US-08-426-5998-2	Sequence 2, Appli
9	7	1.9	207	2 US-08-426-5998-4	Sequence 2, Appli
10	7	1.9	207	3 US-08-813-884-35	Sequence 3, Appli
11	7	1.9	207	3 US-08-813-884-54	Sequence 3, Appli
12	7	1.9	207	6 5221624-1	Patent No. 5221624
13	7	1.9	223	1 US-08-143-579A-4	Sequence 4, Appli
14	7	1.9	223	1 US-08-143-579A-4	Sequence 4, Appli
15	7	1.9	223	1 US-08-454-928-8	Sequence 8, Appli
16	7	1.9	223	2 US-09-154-802-3	Sequence 3, Appli
17	7	1.9	223	3 US-09-373-029-3	Sequence 3, Appli
18	7	1.9	258	3 US-09-053-197A-7	Sequence 7, Appli
19	7	1.9	258	4 US-09-085-761A-7	Sequence 7, Appli
20	7	1.9	901	2 US-08-884-681-5	Sequence 5, Appli
21	7	1.9	901	4 US-09-258-643-5	Sequence 5, Appli
22	7	1.9	942	2 US-08-884-681-4	Sequence 4, Appli
23	7	1.9	942	3 US-08-360-005-6	Sequence 6, Appli
24	7	1.9	942	4 US-09-258-643-4	Sequence 6, Appli
25	7	1.9	942	4 US-09-418-540-6	Sequence 6, Appli
26	7	1.9	968	4 US-08-560-005-7	Sequence 7, Appli
27	7	1.9	968	4 US-09-418-540-7	Sequence 7, Appli

28	6	1.6	8	4 US-08-444-818-462	Sequence 462, App
29	6	1.6	8	4 US-08-444-818-463	Sequence 463, App
30	6	1.6	9	1 US-08-102-738-15	Sequence 15, Appli
31	6	1.6	9	2 US-08-146-028-296	Sequence 296, App
32	6	1.6	9	2 US-08-146-028-297	Sequence 297, App
33	6	1.6	9	2 US-08-146-028-302	Sequence 302, App
34	6	1.6	9	2 US-08-146-028-303	Sequence 303, App
35	6	1.6	9	2 US-08-146-028-304	Sequence 304, App
36	6	1.6	9	2 US-08-146-028-305	Sequence 305, App
37	6	1.6	9	2 US-08-146-028-310	Sequence 310, App
38	6	1.6	9	2 US-08-146-028-311	Sequence 311, App
39	6	1.6	9	2 US-08-845-926-37	Sequence 37, Appli
40	6	1.6	9	2 US-08-765-452-23	Sequence 23, Appli
41	6	1.6	9	4 US-08-723-425A-296	Sequence 296, App
42	6	1.6	9	4 US-08-723-425A-297	Sequence 297, App
43	6	1.6	9	4 US-08-723-425A-302	Sequence 302, App
44	6	1.6	9	4 US-08-723-425A-303	Sequence 303, App
45	6	1.6	9	4 US-08-723-425A-304	Sequence 304, App
46	6	1.6	9	4 US-08-723-425A-305	Sequence 305, App
47	6	1.6	9	4 US-08-723-425A-310	Sequence 310, App
48	6	1.6	9	4 US-08-723-425A-311	Sequence 311, App
49	6	1.6	9	4 US-08-604-365-7	Sequence 7, Appli
50	6	1.6	9	4 US-09-112-206-296	Sequence 296, App
51	6	1.6	9	4 US-09-112-206-297	Sequence 297, App
52	6	1.6	9	4 US-09-112-206-302	Sequence 302, App
53	6	1.6	9	4 US-09-112-206-303	Sequence 303, App
54	6	1.6	9	4 US-09-112-206-304	Sequence 304, App
55	6	1.6	9	4 US-09-112-206-305	Sequence 305, App
56	6	1.6	9	4 US-09-112-206-310	Sequence 310, App
57	6	1.6	9	4 US-09-112-206-311	Sequence 311, App
58	6	1.6	10	3 US-08-386-385-2	Sequence 2, Appli
59	6	1.6	10	4 US-08-287-221-2	Sequence 2, Appli
60	6	1.6	10	4 US-08-938-085A-55	Sequence 55, Appli
61	6	1.6	12	1 US-08-615-279-13	Sequence 13, Appli
62	6	1.6	12	1 US-08-615-279-35	Sequence 35, Appli
63	6	1.6	15	4 US-08-604-365-3	Sequence 3, Appli
64	6	1.6	17	1 US-08-244-1168-10	Sequence 10, Appli
65	6	1.6	17	1 US-08-727-688-34	Sequence 34, Appli
66	6	1.6	18	2 US-08-466-975A-11	Sequence 11, Appli
67	6	1.6	20	2 US-08-466-975A-12	Sequence 12, Appli
68	6	1.6	20	2 US-08-391-671A-11	Sequence 11, Appli
69	6	1.6	20	2 US-08-391-671A-12	Sequence 12, Appli
70	6	1.6	20	2 US-08-853-623D-27	Sequence 27, Appli
71	6	1.6	20	3 US-08-467-902A-11	Sequence 11, Appli
72	6	1.6	20	3 US-08-467-902A-12	Sequence 12, Appli
73	6	1.6	20	4 US-09-275-265-11	Sequence 11, Appli
74	6	1.6	20	4 US-09-275-265-12	Sequence 12, Appli
75	6	1.6	22	2 US-08-146-028-51	Sequence 51, Appli
76	6	1.6	22	2 US-08-146-028-52	Sequence 52, Appli
77	6	1.6	22	2 US-08-146-028-53	Sequence 53, Appli
78	6	1.6	22	2 US-08-146-028-116	Sequence 116, App
79	6	1.6	22	2 US-08-146-028-142	Sequence 142, App
80	6	1.6	22	2 US-08-723-425A-51	Sequence 51, Appli
81	6	1.6	22	4 US-08-723-425A-52	Sequence 52, Appli
82	6	1.6	22	4 US-08-723-425A-53	Sequence 53, Appli
83	6	1.6	22	4 US-08-723-425A-116	Sequence 116, App
84	6	1.6	22	4 US-08-723-425A-142	Sequence 142, App
85	6	1.6	22	4 US-08-723-425A-150	Sequence 150, App
86	6	1.6	22	4 US-09-112-206-51	Sequence 51, Appli
87	6	1.6	22	4 US-09-112-206-52	Sequence 52, Appli
88	6	1.6	22	4 US-09-112-206-53	Sequence 53, Appli
89	6	1.6	22	4 US-09-112-206-142	Sequence 142, App
90	6	1.6	22	4 US-09-112-206-150	Sequence 150, App
91	6	1.6	25	2 US-07-946-054-7	Sequence 7, Appli
92	6	1.6	30	1 US-08-336-553A-17	Sequence 17, Appli
93	6	1.6	30	3 PCT-US93-08638-7	Sequence 7, Appli
94	6	1.6	31	1 US-07-666-719-11	Sequence 11, Appli
95	6	1.6	33	2 US-08-146-028-57	Sequence 57, Appli
96	6	1.6	33	2 US-08-146-028-106	Sequence 106, App

## ALIGNMENTS

```
RESULT 1
US-08-884-681-1
; Sequence 1, Application US/08884681
; Patent No. 5955338
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: NEW HUMAN PHOSPHATIDYLINOSITOL
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/884,681
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0334 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT03
; CLONE: 638789
US-08-884-681-1
```

```
Query Match 100.0%; Score 372; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MDVLSPLSFIVSHVMGILLVFAKYQHLPTIQLSTKSTPTGGLFGYWGKGGVNICL 60
DB 1 MDVLSPLSFIVSHVMGILLVFAKYQHLPTIQLSTKSTPTGGLFGYWGKGGVNICL 60
QY 61 KLYGYVSIINCHLPPHISNNYORLEHFDRILEMQNCGRDIPNILDHLLIIFGDMNR 120
DB 61 KLYGYVSIINCHLPPHISNNYORLEHFDRILEMQNCGRDIPNILDHLLIIFGDMNR 120
QY 121 IEDGGLHVRRESIKKRCYGGIEMKQDLSIAKKHDLREFQEGRIILFPPTKKFDNSNDY 180
DB 121 IEDGGLHVRRESIKKRCYGGIEMKQDLSIAKKHDLREFQEGRIILFPPTKKFDNSNDY 180
QY 181 DTSEKRRKPAWTDRLMLRKQPCAGPDTPIPPASHFSLSLRGYSSHMTYGISDHRKPVSG 240
DB 181 DTSEKRRKPAWTDRLMLRKQPCAGPDTPIPPASHFSLSLRGYSSHMTYGISDHRKPVSG 240
```

```
QY 241 TFELEKPLVSAPLIVLMPEDLMTVENDMVSYSTSDPSSPMDMIGLYKVLADVNDY 300
DB 241 TFELEKPLVSAPLIVLMPEDLMTVENDMVSYSTSDPSSPMDMIGLYKVLADVNDY 300
QY 301 VSIAMVGDGSKVSCSDNLNOVYIDISNPTFEDEFLCTYSNSLRVVISRPFQIPSSL 360
DB 301 VSIAMVGDGSKVSCSDNLNOVYIDISNPTFEDEFLCTYSNSLRVVISRPFQIPSSL 360
QY 361 REDPLGEAQPOT 372
DB 361 REDPLGEAQPOT 372
```

```
RESULT 2
US-09-258-643-1
; Sequence 1, Application US/09258643
; Patent No. 627373
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: NEW HUMAN PHOSPHATIDYLINOSITOL
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/258,643
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/884,681
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0334 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT03
; CLONE: 638789
US-09-258-643-1
```

```
Query Match 100.0%; Score 372; DB 4; Length 372;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MDVLSPLSFIVSHVMGILLVFAKYQHLPTIQLSTKSTPTGGLFGYWGKGGVNICL 60
DB 1 MDVLSPLSFIVSHVMGILLVFAKYQHLPTIQLSTKSTPTGGLFGYWGKGGVNICL 60
QY 61 KLYGYVSIINCHLPPHISNNYORLEHFDRILEMQNCGRDIPNILDHLLIIFGDMNR 120
DB 61 KLYGYVSIINCHLPPHISNNYORLEHFDRILEMQNCGRDIPNILDHLLIIFGDMNR 120
```

QY 121 IEDFGLHFVRESIKRRCYGLMEKQOLSIKKHDPILREFOGRLLPPTYKFDNSNDY 180  
DB 121 IEDFGLHFVRESIKRRCYGLMEKQOLSIKKHDPILREFOGRLLPPTYKFDNSNDY 180  
QY 181 DTSEKKRPAMTDRLIMRLKRPCAGPDTPIPPASHFSLSLRGYSNHTYGISDHRPVSG 240  
DB 181 DTSEKKRPAMTDRLIMRLKRPCAGPDTPIPPASHFSLSLRGYSNHTYGISDHRPVSG 240  
QY 241 TFDLELKVAPLIVLMPEDLMTVENDMVSYSTSDPSSPMDWIGLYKGLRDVNDY 300  
DB 241 TFDLELKVAPLIVLMPEDLMTVENDMVSYSTSDPSSPMDWIGLYKGLRDVNDY 300  
QY 301 VSYAWGSKVSCSDNLNOVYIDISNPTDEFLCYSSLSRVSIGSRFQIPPSGL 360  
DB 301 VSYAWGSKVSCSDNLNOVYIDISNPTDEFLCYSSLSRVSIGSRFQIPPSGL 360  
QY 361 REDPLGEAQPQI 372  
DB 361 REDPLGEAQPQI 372

RESULT 3  
US-08-884-681-3  
Sequence 3, Application US/08884681.  
Patent No. 5955338  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: NEW HUMAN PHOSPHATIDYLINOSITOL  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/884,681  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0334 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 329 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 1399101  
US-08-884-681-3

Query Match 79.3%; Score 295; DB 2; Length 329;  
Best Local Similarity 100.0%; Pred. No. 2.5e-287;  
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLFGYWGKGGVNTICKLITYGYVSTINCHLPHPHISNNYQRLHPRILLEMONECGRDIPN 104  
DB 2 GLFGYWGKGGVNTICKLITYGYVSTINCHLPHPHISNNYQRLHPRILLEMONECGRDIPN 61  
QY 105 ILDHDLITWFDGDMNRIDPFGHVFRESIKRRCYGLMEKQOLSIKKHDPILREFOGR 164  
DB 62 ILDHDLITWFDGDMNRIDPFGHVFRESIKRRCYGLMEKQOLSIKKHDPILREFOGR 121  
QY 165 LFPPTYKFDNSNDYDTSEKKRPAMTDRLIMRLKRPCAGPDTPIPPASHFSLSLRGY 224  
DB 122 LFPPTYKFDNSNDYDTSEKKRPAMTDRLIMRLKRPCAGPDTPIPPASHFSLSLRGY 181  
QY 225 SSHMTYGISDHRKPVSGTDELKPLVAPLIVLMPEDLMTVENDMVSYSTSDPSSPW 284  
DB 182 SSHMTYGISDHRKPVSGTDELKPLVAPLIVLMPEDLMTVENDMVSYSTSDPSSPW 241  
QY 285 DWIGLYKGLRDVNDYVSYAWGSKVSCSDNLNOVYIDISNPTDEFLCY 339  
DB 242 DWIGLYKGLRDVNDYVSYAWGSKVSCSDNLNOVYIDISNPTDEFLCY 296

RESULT 4  
US-09-258-643-3  
Sequence 3, Application US/09258643  
Patent No. 627373  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: NEW HUMAN PHOSPHATIDYLINOSITOL  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/258,643  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/884,681  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0334 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 329 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 1399101  
US-09-258-643-3

Query Match 79.3%; Score 295; DB 4; Length 329;  
Best Local Similarity 100.0%; Pred. No. 2.5e-287;  
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLFGYGNKGGVNICLKLYGYVSIINCHLPPHISNNYORLEHEDRILEMONECEGRDIPN 104  
|||||  
Db 2 GLFGYGNKGGVNICLKLYGYVSIINCHLPPHISNNYORLEHEDRILEMONECEGRDIPN 61  
QY 105 ILDDHLLIWFEDMNRIEDFGLHFVRESIKKRCYGGIEMEKOLSTAKKHDDLRFQGR 164  
|||||  
Db 62 ILDDHLLIWFEDMNRIEDFGLHFVRESIKKRCYGGIEMEKOLSTAKKHDDLRFQGR 121  
QY 165 LLFPPTYKFDNSNDYDTSEKKRRKPAWDRILMLRKPACGPDPIPPASHFSLRGY 224  
|||||  
Db 122 LLFPPTYKFDNSNDYDTSEKKRRKPAWDRILMLRKPACGPDPIPPASHFSLRGY 181  
QY 225 SSHWTGYSIDHKPVSQGTDELKPLVSAPLIVLPEDLMTVENDMVSYSTDSFPSSPW 284  
|||||  
Db 182 SSHWTGYSIDHKPVSQGTDELKPLVSAPLIVLPEDLMTVENDMVSYSTDSFPSSPW 241  
QY 285 DWIGLYKGLRDVNDYVYAWVGDSKSCDNLNOVYIDISNPTTEDEFLICY 339  
|||||  
Db 242 DWIGLYKGLRDVNDYVYAWVGDSKSCDNLNOVYIDISNPTTEDEFLICY 296  
RESULT 5  
US-08-560-005-4  
; Sequence 4, Application US/08560005  
; Patent No. 6001354  
; GENERAL INFORMATION:  
; APPLICANT: Pot, David A.  
; APPLICANT: Williams, Lewis T.  
; APPLICANT: Jefferson, Anne Bennett  
; APPLICANT: Majerus, Philip W.  
; TITLE OF INVENTION: No. 6001354e1 Grb2 Associating Protein and Nucleic  
; TITLE OF INVENTION: Acids Encoding Therefor  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew  
; STREET: One Market Plaza, Stewart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/560,005  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 2307K-0624000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 946 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..946  
; OTHER INFORMATION: /note= "ysc5ptase"  
US-08-560-005-4

Query Match 2.2%; Score 8; DB 3; Length 946;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 FPPYTKFD 174  
|||||  
Db 762 FPPYTKFD 769  
RESULT 6  
US-09-418-540-4  
; Sequence 4, Application US/09418540  
; Patent No. 6296848  
; GENERAL INFORMATION:  
; APPLICANT: Pot, David A.  
; APPLICANT: Williams, Lewis T.  
; APPLICANT: Jefferson, Anne Bennett  
; APPLICANT: Majerus, Philip W.  
; TITLE OF INVENTION: No. 6296848e1 Grb2 Associating Protein and Nucleic  
; TITLE OF INVENTION: Acids Encoding Therefor  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew  
; STREET: One Market Plaza, Stewart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/418,540  
; FILING DATE: 14-OCT-1999  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/560,005  
; FILING DATE: 17-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 2307K-0624000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 946 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..946  
; OTHER INFORMATION: /note= "ysc5ptase"  
US-09-418-540-4

Query Match 2.2%; Score 8; DB 4; Length 946;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 FPPYTKFD 174  
|||||  
Db 762 FPPYTKFD 769

RESULT 7  
5221624-30  
; Patent No. 5221624  
; APPLICANT: BLAIR, LINDLEY C.;KODURI, JAR-HOW;WEICKMANN,  
; JOACHIM J.  
; TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND  
; (LYS46, ASP113, ASP137) THAUMATIN I  
; NUMBER OF SEQUENCES: 31

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/407,416  
FILING DATE: 14-SEP-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 268,702  
FILING DATE: 08-NOV-1988  
SEQ ID NO:30:  
LENGTH: 206  
5221624-30

## Query Match

Best Local Similarity 1.9%; Score 7; DB 6; Length 206;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 321 YIDISNI 327  
|||||  
Db 99 YIDISNI 105

RESULT 8  
US-08-426-599B-2

Sequence 2, Application US/08426599B  
Patent No. 5932438

## GENERAL INFORMATION:

APPLICANT: Uriach-Marsal, Juan  
APPLICANT: Rubio-Susan, Victor  
APPLICANT: Palino-Martin, Cristina  
APPLICANT: Kalo-Koenova, Eliza I.  
APPLICANT: del Moral-Juarez, Catalina  
APPLICANT: Faus-Santasusana, Ignacio  
APPLICANT: del Rio-Pericacho, Jose-Luis  
TITLE OF INVENTION: Preparation of Thaumatin Sweeteners  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
STREET: 555 Thirteenth Street, N.W., Suite 701 East  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/426,599B  
FILING DATE: 21-APR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Saxe, Stephen A.  
REGISTRATION NUMBER: 38,609  
REFERENCE/DOCKET NUMBER: 1604-123A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-624-1589  
TELEFAX: 202-624-6031  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-426-599B-2

US-08-426-599B-2

## Query Match

Best Local Similarity 1.9%; Score 7; DB 2; Length 207;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 321 YIDISNI 327

Db 99 YIDISNI 105

RESULT 9

US-08-426-599B-4  
Sequence 4, Application US/08426599B  
Patent No. 5932438

## GENERAL INFORMATION:

APPLICANT: Uriach-Marsal, Juan  
APPLICANT: Rubio-Susan, Victor  
APPLICANT: Palino-Martin, Cristina  
APPLICANT: Kalo-Koenova, Eliza I.  
APPLICANT: del Moral-Juarez, Catalina  
APPLICANT: Faus-Santasusana, Ignacio  
APPLICANT: del Rio-Pericacho, Jose-Luis  
TITLE OF INVENTION: Preparation of Thaumatin Sweeteners  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
STREET: 555 Thirteenth Street, N.W., Suite 701 East  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/426,599B  
FILING DATE: 21-APR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Saxe, Stephen A.  
REGISTRATION NUMBER: 38,609  
REFERENCE/DOCKET NUMBER: 1604-123A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-624-1589  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-426-599B-4

## Query Match

Best Local Similarity 1.9%; Score 7; DB 2; Length 207;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 321 YIDISNI 327  
|||||  
Db 99 YIDISNI 105

RESULT 10

US-08-813-884-35  
Sequence 35, Application US/08813884C  
Patent No. 6001410

## GENERAL INFORMATION:

APPLICANT: Bolen, Paul L.  
APPLICANT: Kosiakoff, Nicholas  
APPLICANT: Hawn, Regina  
APPLICANT: Scharpf Jr., Lewis G.  
TITLE OF INVENTION: A FRUIT LIQUOR BEVERAGE CONTAINING MONELLIN  
FILE REFERENCE: A fruit liqueur beverage...SCM...enhan

;; CURRENT APPLICATION NUMBER: US/08/813,884C  
;; CURRENT FILING DATE: 1997-03-06  
;; EARLIER APPLICATION NUMBER: 60/022,597  
;; EARLIER FILING DATE: 1996-07-25  
;; NUMBER OF SEQ ID NOS: 54  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 35  
;; LENGTH: 207  
;; TYPE: PRF  
;; ORGANISM: Thaumatooccus daniellii  
;; PUBLICATION INFORMATION:  
;; TITLE: DNA ENCODING (LYS46, ASP97, ASP113) AND (LYS46, ASP113,  
;; DATABASE ACCESSION NUMBER: US005221624A  
;; PATENT DOCUMENT NUMBER: US 5,221,624  
;; PATENT FILING DATE: 1989-09-14  
;; PUBLICATION DATE: 1993-06-22  
US-08-813-884-35

Query Match 1.9%; Score 7; DB 3; Length 207;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 321 YIDISNI 327  
|||||  
DB 99 YIDISNI 105

RESULT 11  
US-08-813-884-54  
;; Sequence 54, Application US/08813884C  
;; Patent No. 6001410  
;; GENERAL INFORMATION:  
;; APPLICANT: Bolen, Paul L  
;; APPLICANT: Kossiakoff, Nicholas  
;; APPLICANT: Hawn, Regina  
;; APPLICANT: Scharp, J., Lewis G.  
;; TITLE OF INVENTION: A FRUIT LIQUOR BEVERAGE CONTAINING RECOMBINANT MONELLIN  
;; TITLE OF INVENTION: TO ENHANCE THE ALCOHOLIC IMPACT  
;; FILE REFERENCE: A fruit liquor beverage...SCM...enhan  
;; CURRENT APPLICATION NUMBER: US/08/813,884C  
;; CURRENT FILING DATE: 1997-03-06  
;; EARLIER APPLICATION NUMBER: 60/022,597  
;; EARLIER FILING DATE: 1996-07-25  
;; NUMBER OF SEQ ID NOS: 54  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 54  
;; LENGTH: 207  
;; TYPE: PRF  
;; ORGANISM: Thaumatooccus daniellii  
;; PUBLICATION INFORMATION:  
;; TITLE: DNA ENCODING (LYS46, ASP97, ASP113) AND (LYS46, ASP113,  
;; DATABASE ACCESSION NUMBER: US005221624A  
;; PATENT DOCUMENT NUMBER: US 5,221,624  
;; PATENT FILING DATE: 1989-09-14  
;; PUBLICATION DATE: 1993-06-22  
US-08-813-884-54

Query Match 1.9%; Score 7; DB 3; Length 207;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 321 YIDISNI 327  
|||||  
DB 99 YIDISNI 105

RESULT 12  
5221624-1  
;; Patent No. 5221624

;; APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,  
;; JOACHIM J.  
;; TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND  
;; (LYS46, ASP113, ASP137) THAUMATIN I  
;; NUMBER OF SEQUENCES: 31  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/407,416  
;; FILING DATE: 14-SEP-1989  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 268,702  
;; FILING DATE: 08-NOV-1988  
;; SEQ ID NO.1:  
;; LENGTH: 207  
5221624-1

Query Match 1.9%; Score 7; DB 6; Length 207;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 321 YIDISNI 327  
|||||  
DB 99 YIDISNI 105

RESULT 13  
US-08-143-579A-4  
;; Sequence 4, Application US/08143579A  
;; Patent No. 5625034  
;; GENERAL INFORMATION:  
;; APPLICANT: Liao, Jaw-Ching  
;; TITLE OF INVENTION: CORE ANTIGEN PROTEIN OF HEPATITIS C  
;; TITLE OF INVENTION: VIRUS, AND DIAGNOSTIC METHOD AND KIT USING THE SAME  
;; NUMBER OF SEQUENCES: 4  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Seed and Berry  
;; STREET: 6300 Columbia Center, 701 Fifth Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: U.S.  
;; ZIP: 98104-7092  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/143,579A  
;; FILING DATE: 26-OCT-1993  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963483  
;; FILING DATE: 16-OCT-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: King, Joshua  
;; REGISTRATION NUMBER: 35,570  
;; REFERENCE/DOCKET NUMBER: 310068,402D1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206) 622-4900  
;; TELEFAX: (206) 682-6031  
;; TELEX: 3723836  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 223 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
US-08-143-579A-4

Query Match 1.9%; Score 7; DB 1; Length 223;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 09:24:59 ; Search time 21.78 Seconds

(without alignments)  
1641.195 Million cell updates/sec

Title: US-09-892-287-1

Perfect score: 372

Sequence: 1 MDVSPLSFIVSHVHMQGT.....FQIPPGSLREPLGEAPQI 372

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Wc size: 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : PIR\_71:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.4	427	2	C88883
2	8	2.2	262	2	A81504
3	8	2.2	262	2	B72037
4	8	2.2	262	2	E86588
5	8	2.2	331	2	D96515
6	8	2.2	713	2	A35502
7	8	2.2	753	2	T19338
8	8	2.2	869	1	A28443
9	8	2.2	946	2	S48433
10	8	2.2	1076	2	T40141
11	8	2.2	1332	2	D84669
12	7	1.9	143	2	AC2693
13	7	1.9	146	2	G97474
14	7	1.9	154	2	S30728
15	7	1.9	179	2	A84684
16	7	1.9	188	2	F87217
17	7	1.9	193	2	A81208
18	7	1.9	193	2	B81785
19	7	1.9	207	1	QTTCL
20	7	1.9	235	1	QTTCL
21	7	1.9	235	1	H75539
22	7	1.9	255	2	S51502
23	7	1.9	258	2	H65188
24	7	1.9	258	2	H91224
25	7	1.9	258	2	F86071
26	7	1.9	259	2	AG0459
27	7	1.9	259	2	T09317
28	7	1.9	276	2	H71147
29	7	1.9	280	2	T09939

30	7	1.9	285	2	B83047	hypothetical prote
31	7	1.9	289	2	S04648	ampr protein - Rho
32	7	1.9	294	2	S38983	kinesin-related pr
33	7	1.9	317	2	H82785	dolichol-phosphate
34	7	1.9	317	2	AF0205	probable exported
35	7	1.9	337	2	S38166	hypothetical prote
36	7	1.9	342	2	D84180	hypothetical prote
37	7	1.9	350	2	G95009	sensor histidine k
38	7	1.9	350	2	E97881	probable exported
39	7	1.9	381	2	AG0110	probable membrane
40	7	1.9	419	2	T33384	hypothetical prote
41	7	1.9	458	2	B81409	probable transmemb
42	7	1.9	462	2	T25544	hypothetical prote
43	7	1.9	474	2	B64434	anthranilate synth
44	7	1.9	483	2	G87493	conserved hypothet
45	7	1.9	484	2	H83084	glu-tRNA(Gln) amid
46	7	1.9	520	2	A37844	hypothetical prote
47	7	1.9	526	2	A56573	nuclear pore compl
48	7	1.9	552	2	A64085	probable ATP-bindi
49	7	1.9	578	2	B64012	hypothetical prote
50	7	1.9	615	2	AH2028	hypothetical prote
51	7	1.9	641	2	D83046	selenocysteine-spe
52	7	1.9	666	2	T20997	hypothetical prote
53	7	1.9	672	2	A41075	inositol-1,4,5-tri
54	7	1.9	742	1	S58691	kinesin-related pr
55	7	1.9	747	1	A57107	kinesin-related pr
56	7	1.9	793	2	JC5831	kinesin-related pr
57	7	1.9	794	2	T40289	hypothetical prote
58	7	1.9	813	2	G96494	protein F7F22.4 (f
59	7	1.9	842	2	B55853	outer membrane ush
60	7	1.9	880	2	T00800	disease resistance
61	7	1.9	905	2	T40015	phosphatidylethano
62	7	1.9	935	2	T39389	probable DNA damag
63	7	1.9	970	2	S29069	inositol polyphosp
64	7	1.9	993	2	G96680	F514.11 [imported
65	7	1.9	1017	2	T42384	inositol-1,4,5-tri
66	7	1.9	1130	2	T21134	hypothetical prote
67	7	1.9	1136	2	D86190	hypothetical prote
68	7	1.9	1183	2	F84725	probable membrane
69	7	1.9	1184	2	S63046	probable inositol
70	7	1.9	1305	2	T00670	probable inositol
71	7	1.9	1314	2	T09481	matng type silenc
72	7	1.9	1371	2	T42938	major capsid prote
73	7	1.9	1533	2	T00344	hypothetical prote
74	7	1.9	1541	2	S46686	hypothetical prote
75	7	1.9	1575	2	S68448	synaptotjanin, 170K
76	7	1.9	3149	1	QOBE8	BLP1 protein - hu
77	7	1.6	32	2	D60580	growth hormone-rel
78	6	1.6	35	2	E95098	hypothetical prote
79	6	1.6	36	2	D82466	hypothetical prote
80	6	1.6	45	2	JH0208	hypothetical prote
81	6	1.6	53	2	D72610	hypothetical prote
82	6	1.6	54	2	D31443	ovomucoid, third d
83	6	1.6	60	2	S02189	hypothetical prote
84	6	1.6	62	2	H81791	hypothetical prote
85	6	1.6	62	2	C97847	hypothetical prote
86	6	1.6	74	2	S24965	probable proteinas
87	6	1.6	75	2	S51637	probable proteinas
88	6	1.6	78	2	D86654	oxidoreductase ycd
89	6	1.6	82	2	E70090	hypothetical prote
90	6	1.6	82	2	S07352	hypothetical prote
91	6	1.6	85	2	E83280	hypothetical prote
92	6	1.6	89	2	F82521	hypothetical prote
93	6	1.6	90	2	S75278	hypothetical prote
94	6	1.6	94	2	A86739	hypothetical prote
95	6	1.6	95	2	B81064	hypothetical prote
96	6	1.6	101	2	T02694	hypothetical prote
97	6	1.6	102	2	H25035	hypothetical prote
98	6	1.6	102	2	G25035	hypothetical prote
99	6	1.6	105	2	G97090	chlorodoxin, ltr l
100	6	1.6				

## ALIGNMENTS

RESULT 1  
C88883  
protein JC8.10 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: C88883  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A:Reference number: A75000; MID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_elg  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: C88883  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-427 <STO>  
A:Cross-references: GB:chr\_IV; PIDN:CA805234.1; PID:g3924785; GSPDB:GN00022; CESP:JC8.10  
A:Gene: JC8.10  
A:Map position: 4

Query Match 2.4%; Score 8; DB 2; Length 427;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 FAWTDRIILM 197  
DB 102 FAWTDRIILM 110

RESULT 2  
A81504  
conserved hypothetical protein CP1089 [imported] - Chlamydia pneumoniae (strain AR39  
C:Species: Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C:Accession: A81504  
R:Read, T.D.; Brumham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MID:20150255  
A:Accession: A81504  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-262 <REA>  
A:Cross-references: GB:AE002265; GB:AE002161; MID:g7189999; PIDN:AAF38860.1; PID:g719000  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP1089

Query Match 2.2%; Score 8; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GILLVFA 26  
DB 17 GILLVFA 24

RESULT 3  
B72037  
ct598 hypothetical protein - Chlamydia pneumoniae (strain CML029)  
C:Species: Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 05-May-2000  
C:Accession: B72037  
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MID:99206606  
A:Accession: B72037  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-262 <ARN>  
A:Cross-references: GB:AE001659; GB:AE001363; MID:g4377081; PIDN:AAD18921.1; PID:g437  
A:Experimental source: strain CML029  
C:Genetics:  
A:Gene: CPn0783

Query Match 2.2%; Score 8; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GILLVFA 26  
DB 17 GILLVFA 24

RESULT 4  
E86588  
CT598 hypothetical protein [imported] - Chlamydia pneumoniae (strain J138)  
C:Species: Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
C:Accession: E86588  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A:Reference number: A86491; MID:20330349  
A:Accession: E86588  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-262 <STO>  
A:Cross-references: GB:BA000008; MID:g8979156; PIDN:BA89891.1; GSPDB:GN00142  
C:Genetics:  
A:Experimental source: strain J138  
A:Gene: CPj0783

Query Match 2.2%; Score 8; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GILLVFA 26  
DB 17 GILLVFA 24

RESULT 5  
D96515  
hypothetical protein F16N3.22 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: D96515  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marita  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MID:21016719  
A:Accession: D96515  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-331 <STO>  
A:Cross-references: GB:AE005173; MID:g5668810; PIDN:AAD46036.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F16N3.22



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 09:27:29 ; Search time 12.29 Seconds

(without alignments)  
1171.983 Million cell updates/sec

Title: US-09-892-287-1

Perfect score: 372

Sequence: 1 MDVLSPLSFIVSHVMOGL.....FOIPPGSLRDEPLGAPQPI 372

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

W size: 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	2.2	713	1 TSA4_GIALA	P21849 giardia lam
2	8	2.2	869	1 PEM1_YEAST	P05374 saccharomyc
3	8	2.2	946	1 YIA2_YEAST	P40559 saccharomyc
4	7	1.9	207	1 THM1_THADA	P02883 thaumatococ
5	7	1.9	235	1 THM2_THADA	P02884 thaumatococ
6	7	1.9	255	1 ATP6_YARLI	Q36258 yarrowia li
7	7	1.9	258	1 TATC_ECOLI	P27857 escherichia
8	7	1.9	289	1 AMPR_RHOCA	P14145 rhodobacter
9	7	1.9	337	1 YK68_YEAST	P36164 saccharomyc
10	7	1.9	474	1 TRPE_METJA	O58475 methanococc
11	7	1.9	478	1 CARL_HUMAN	O94076 homo sapien
12	7	1.9	520	1 BAIB_EUBSP	P19409 eubacterium
13	7	1.9	552	1 Y664_HAEIN	O57538 haemophilus
14	7	1.9	742	1 YTFM_HAEIN	P44038 haemophilus
15	7	1.9	742	1 KIZ1_STRPU	P46871 strongyloce
16	7	1.9	747	1 KF3B_HUMAN	O15066 homo sapien
17	7	1.9	747	1 KF3B_MOUSE	O61771 mus musculu
18	7	1.9	793	1 KF3C_HUMAN	O14782 homo sapien
19	7	1.9	794	1 YGVY_SCHPO	O43033 schizosacch
20	7	1.9	796	1 KF3C_MOUSE	O35066 mus musculu
21	7	1.9	796	1 KF3C_MOUSE	O55165 rattus norv
22	7	1.9	842	1 AGGC_ECOLI	P46005 escherichia
23	7	1.9	901	1 OCRI_HUMAN	O01968 homo sapien
24	7	1.9	905	1 CHO2_SCHPO	O74787 schizosacch
25	7	1.9	942	1 ISP2_HUMAN	P32019 homo sapien
26	7	1.9	1183	1 YNKG_YEAST	P50942 saccharomyc
27	7	1.9	1239	1 DPG1_MOUSE	P54099 mus musculu
28	7	1.9	1334	1 SYU1_BOVIN	O18964 bos taurus
29	7	1.9	1333	1 Y586_HUMAN	O9bvv6 homo sapien
30	7	1.9	1541	1 YH16_YEAST	P38873 saccharomyc
31	7	1.9	1547	1 TOP2_BOMMO	O16140 bombyx mori
32	7	1.9	1574	1 SYU1_RAT	O62910 rattus norv
33	7	1.9	1575	1 SYU1_HUMAN	O43426 homo sapien

34	7	1.9	3149	1 TEGU_EBV	P03186 epstein-bar
35	6	1.6	32	1 GHR4_RAT	P33581 rattus norv
36	6	1.6	54	1 IOVO_LEIOC	P05581 leiopoa ocel
37	6	1.6	54	1 YE3B_HAEIN	O86239 haemophilus
38	6	1.6	60	1 PHX1_MOUSE	P15973 mus musculu
39	6	1.6	74	1 SRP_SOYBN	O07502 glycine max
40	6	1.6	82	1 YCX1_VICFA	P08889 vicia faba
41	6	1.6	108	1 YG5N_YEAST	P40326 saccharomyc
42	6	1.6	125	1 IFM1_HUMAN	P13164 homo sapien
43	6	1.6	139	1 ATP6_DICDH	P30159 dicytota di
44	6	1.6	143	1 MIOC_PHOLA	P20094 photoridium
45	6	1.6	146	1 MIOC_ECO57	P58208 escherichia
46	6	1.6	146	1 MIOC_ECOLI	P03817 escherichia
47	6	1.6	153	1 SODC_DESHA	O42724 debaryomyce
48	6	1.6	153	1 YLR3_EBVP3	O07286 epstein-bar
49	6	1.6	154	1 SODC_SCHPO	P28758 schizosacch
50	6	1.6	161	1 VANZ_ENTFC	O06242 enterococc
51	6	1.6	176	1 IL17_SHEEP	O28540 ovis aries
52	6	1.6	177	1 IL17_HUMAN	P13232 homo sapien
53	6	1.6	179	1 ARF1_SCHPO	P13579 schizosacch
54	6	1.6	181	1 ARF_CRYNE	P34728 cryptococcus
55	6	1.6	182	1 ARF_AJECA	P34727 ajellomyces
56	6	1.6	182	1 WCAE_ECOLI	P71240 escherichia
57	6	1.6	183	1 ATPD_BACHD	O9k6h2 bacillus ha
58	6	1.6	184	1 ADML_MOUSE	P97287 mus musculu
59	6	1.6	186	1 MAUE_PARDE	P29886 paracoccus
60	6	1.6	191	1 SPC3_CANFA	P13679 canis fami
61	6	1.6	191	1 SPC3_HUMAN	O9bvy0 homo sapien
62	6	1.6	191	1 SPC3_MOUSE	O9d8v7 mus musculu
63	6	1.6	191	1 SPC3_RAT	O9wrt7 rattus norv
64	6	1.6	202	1 RS2_CRICR	P46791 cricetus
65	6	1.6	204	1 TNE6_HUMAN	O95857 homo sapien
66	6	1.6	209	1 RRMJ_ECOLI	P28652 escherichia
67	6	1.6	211	1 YOR3_SOOV3	O04550 southampton
68	6	1.6	214	1 PYRF_THEAC	O74110 thermoplasma
69	6	1.6	218	1 BIOD_HELPY	O24872 helicobacte
70	6	1.6	219	1 FLAM_METVO	P27802 methanococc
71	6	1.6	220	1 MAUM_METEX	O49130 methylobact
72	6	1.6	221	1 GTR1_ANTST	P80894 antechinus
73	6	1.6	222	1 GTA3_RAT	P14942 rattus norv
74	6	1.6	222	1 GTA4_MOUSE	P24472 mus musculu
75	6	1.6	235	1 FARE_METJA	O58896 methanococc
76	6	1.6	235	1 YBER_ECOLI	P77637 escherichia
77	6	1.6	236	1 UNG_CHLPN	O92733 chlamydia p
78	6	1.6	237	1 MTRM_MERTH	O27227 methanobact
79	6	1.6	240	1 MTRM_MERTH	P80184 methanobact
80	6	1.6	242	1 UNG_MYCPN	P75536 mycoplasma
81	6	1.6	242	1 GRPE_HALME	O9h5c2 halobacteri
82	6	1.6	246	1 Y171_AOUAF	O66558 aquifex aeo
83	6	1.6	248	1 TPIS_CLOAB	O52633 clostridium
84	6	1.6	248	1 YJHP_ECOLI	P29387 escherichia
85	6	1.6	253	1 T4S8_HUMAN	O60637 homo sapien
86	6	1.6	253	1 T4S8_MOUSE	O9qy3 mus musculu
87	6	1.6	255	1 TATC_AZOC	P54085 azotobacter
88	6	1.6	257	1 YZGI_CAREL	P55376 caenorhabdi
89	6	1.6	259	1 PSA3_CAREL	O09583 caenorhabdi
90	6	1.6	261	1 US13_HCVNA	P09720 humana cytom
91	6	1.6	264	1 GP3D_CHLPS	O46261 chlamydia p
92	6	1.6	265	1 YJ15_SCHPO	O13679 schizosacch
93	6	1.6	265	1 RS2_DICDI	P16659 dictyostell
94	6	1.6	265	1 RS2_LETAM	O43992 leishmania
95	6	1.6	266	1 IF2A_AERPE	O9y102 aeropyrum p
96	6	1.6	267	1 RS2_DROME	P31009 drosophila
97	6	1.6	270	1 YAEI_ECOLI	P31704 escherichia
98	6	1.6	272	1 RS2_CAREL	P51403 caenorhabdi
99	6	1.6	274	1 PYRE_MYCTU	P77898 mycobacteri
100	6	1.6	276	1 YGRF_ECOLI	O46818 escherichia

## ALIGNMENTS

RESULT 1



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 09:28:04 ; Search time 28.11 Seconds  
(without alignments)  
2289.367 Million cell updates/sec

Title: US-09-892-287-1

Perfect score: 372  
Sequence: 1 MVLSPSLRISIKVSHVMQGI.....FQIPPSLRKEDPLGEAQPQI 372

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 562222 segs, 172994929 residues

W size: 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

1: SPREMBL\_19:\*  
2: SP\_bacteria:\*  
3: SP\_fungi:\*  
4: SP\_human:\*  
5: SP\_invertebrate:\*  
6: SP\_mammal:\*  
7: SP\_mhc:\*  
8: SP\_organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_rodent:\*  
12: SP\_virus:\*  
13: SP\_vertebrate:\*  
14: SP\_unclassified:\*  
15: SP\_virus:\*  
16: SP\_bacteriaph:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB#	ID	Description
1	372	100.0	448	4	Q9BT40	Q9BT40 homo sapien
2	339	91.1	372	4	Q9P2R5	Q9P2R5 homo sapien
3	339	91.1	448	4	Q9NPJ5	Q9NPJ5 homo sapien
4	295	79.3	329	4	Q157J3	Q157J3 homo sapien
5	24	6.5	468	11	Q09040	Q09040 mus musculu
6	13	3.5	397	4	Q157J5	Q157J5 mus musculu
7	13	3.5	1001	11	Q9JMC1	Q9JMC1 rattus norv
8	13	3.5	1056	4	Q9UDT9	Q9UDT9 homo sapien
9	9	2.4	1113	5	Q9XUD3	Q9XUD3 caenorhabdi
10	9	2.4	1119	5	Q9GT42	Q9GT42 caenorhabdi
11	8	2.2	262	16	Q9Z7C3	Q9Z7C3 chlamydia p
12	8	2.2	262	16	Q9JRY0	Q9JRY0 chlamydia p
13	8	2.2	331	10	Q9SX81	Q9SX81 arabidopsis
14	8	2.2	709	5	Q97444	Q97444 giardia lam
15	8	2.2	753	5	Q17590	Q17590 caenorhabdi
16	8	2.2	1076	3	Q43001	Q43001 schizosacch

17	8	2.2	1171	3	Q96U77	Q96U77 neurospora
18	8	1.9	1332	10	Q9SIC7	Q9SIC7 arabidopsis
19	7	1.9	98	8	Q9G297	Q9G297 eulemur ful
20	7	1.9	98	8	Q9BIB3	Q9BIB3 propithecus
21	7	1.9	98	8	Q9B1I7	Q9B1I7 propithecus
22	7	1.9	98	8	Q9G230	Q9G230 eulemur mac
23	7	1.9	98	8	Q9G9Z8	Q9G9Z8 eulemur ful
24	7	1.9	98	8	Q9G9Z4	Q9G9Z4 eulemur ful
25	7	1.9	98	8	Q9B8T9	Q9B8T9 avahi lanig
26	7	1.9	98	8	Q9B8T5	Q9B8T5 propithecus
27	7	1.9	98	8	Q9B8S7	Q9B8S7 propithecus
28	7	1.9	98	8	Q9G229	Q9G229 eulemur ful
29	7	1.9	98	8	Q9S5U2	Q9S5U2 cheirogaleu
30	7	1.9	98	8	Q9S5T0	Q9S5T0 microcebus
31	7	1.9	98	8	Q9S5S4	Q9S5S4 microcebus
32	7	1.9	98	8	Q9S5S2	Q9S5S2 microcebus
33	7	1.9	98	8	Q94Y87	Q94Y87 lemur catla
34	7	1.9	98	8	Q94XZ5	Q94XZ5 eulemur cor
35	7	1.9	98	8	Q94XZ1	Q94XZ1 eulemur rub
36	7	1.9	98	8	Q94Q41	Q94Q41 microcebus
37	7	1.9	98	8	Q94Q33	Q94Q33 microcebus
38	7	1.9	98	8	Q94Q24	Q94Q24 cheirogaleu
39	7	1.9	98	8	Q94PW7	Q94PW7 cheirogaleu
40	7	1.9	98	8	Q94PA4	Q94PA4 mirza coque
41	7	1.9	98	8	Q94N32	Q94N32 microcebus
42	7	1.9	98	8	Q94N31	Q94N31 microcebus
43	7	1.9	145	2	P70734	P70734 acinetobact
44	7	1.9	164	3	Q9V7B4	Q9V7B4 armillaria
45	7	1.9	165	11	Q9D6I8	Q9D6I8 mus musculu
46	7	1.9	171	2	Q9KX05	Q9KX05 streptomyce
47	7	1.9	179	10	Q9SKN3	Q9SKN3 arabidopsis
48	7	1.9	188	16	Q9CB37	Q9CB37 mycobacteri
49	7	1.9	193	16	Q9K136	Q9K136 neisseria m
50	7	1.9	193	16	Q9JSU4	Q9JSU4 neisseria m
51	7	1.9	204	10	Q9SOJ7	Q9SOJ7 populus tre
52	7	1.9	235	16	Q9RXM6	Q9RXM6 deinococcus
53	7	1.9	236	12	Q69558	Q69558 human herpe
54	7	1.9	235	8	Q9B6D9	Q9B6D9 yarrowia li
55	7	1.9	259	12	Q69047	Q69047 human herpe
56	7	1.9	269	2	Q9X6Z7	Q9X6Z7 bacillus th
57	7	1.9	276	17	Q58131	Q58131 pyrococcus
58	7	1.9	280	10	Q9SU75	Q9SU75 arabidopsis
59	7	1.9	285	16	Q9HV20	Q9HV20 pseudomonas
60	7	1.9	317	16	Q9PEP5	Q9PEP5 xylella las
61	7	1.9	335	4	Q9NKR1	Q9NKR1 homo sapien
62	7	1.9	337	5	Q26681	Q26681 trypanosoma
63	7	1.9	344	11	Q9CZ46	Q9CZ46 halobacteri
64	7	1.9	344	11	Q9CZ46	Q9CZ46 mus musculu
65	7	1.9	350	16	Q9ES39	Q9ES39 rattus norv
66	7	1.9	350	16	Q9S1J1	Q9S1J1 streptococ
67	7	1.9	357	5	Q9V7X0	Q9V7X0 drosophila
68	7	1.9	384	3	Q08227	Q08227 saccharomyc
69	7	1.9	390	16	Q98CB9	Q98CB9 rhizobium l
70	7	1.9	419	5	Q17247	Q17247 caenorhabdi
71	7	1.9	423	3	Q74282	Q74282 coprinus ci
72	7	1.9	428	5	Q9VJ61	Q9VJ61 drosophila
73	7	1.9	433	3	Q9UVN2	Q9UVN2 coprinus ci
74	7	1.9	458	16	Q9PH05	Q9PH05 campylobact
75	7	1.9	462	5	Q02134	Q02134 caenorhabdi
76	7	1.9	483	16	Q9A6W0	Q9A6W0 caulobacter
77	7	1.9	484	16	Q9HW78	Q9HW78 pseudomonas
78	7	1.9	508	8	Q9VXE7	Q9VXE7 drosophila
79	7	1.9	526	11	Q99JN7	Q99JN7 mus musculu
80	7	1.9	526	11	Q63850	Q63850 mus sp. nuc
81	7	1.9	533	10	Q9LU45	Q9LU45 arabidopsis
82	7	1.9	534	3	Q9HE07	Q9HE07 schizosacch
83	7	1.9	607	3	Q9U5H3	Q9U5H3 caenorhabdi
84	7	1.9	612	2	Q69801	Q69801 streptomyce
85	7	1.9	641	16	Q9HW02	Q9HW02 pseudomonas
86	7	1.9	643	5	Q9X222	Q9X222 drosophila
87	7	1.9	666	5	Q19486	Q19486 caenorhabdi
88	7	1.9	744	13	Q93478	Q93478 xenopus lae
89	7	1.9	749	11	Q9DBW2	Q9DBW2 mus musculu

90 7 1.9 776 11 054996 054996 mus musculus  
91 7 1.9 782 5 096V92 096V92 caenorhabditis  
92 7 1.9 782 5 019633 019633 caenorhabditis  
93 7 1.9 794 3 043023 043023 schizosaccharomyces  
94 7 1.9 813 10 09SHN6 09SHN6 arabidopsis  
95 7 1.9 825 11 091ZF8 091ZF8 mus musculus  
96 7 1.9 856 10 0942D7 0942D7 oryza sativa  
97 7 1.9 890 10 048849 048849 arabidopsis  
98 7 1.9 935 3 094623 094623 schizosaccharomyces  
99 7 1.9 950 2 09RJ70 09RJ70 streptomyces  
100 7 1.9 993 10 004475 004475 arabidopsis

## ALIGNMENTS

RESULT 1  
ID 09BT40 PRELIMINARY; PRT: 448 AA.  
09BT40: 01-JUN-2001 (TREMBLrel. 17, Created)  
01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE SKIP FOR SKELETAL MUSCLE AND KIDNEY ENRICHED INOSITOL  
PHOSPHATASE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
NC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=EYE, RETINOBLASTOMA;  
RA Strausberg R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC004362; AAH04362.1; -  
DR InterPro: IPR000300; IPPC:  
DR Pfam: PF00783; IPPC: 1.  
DR SMART: SM00128; IPPC: 1.  
SQ SEQUENCE 448 AA; 51090 MW; 46FAA48C6E2EAD4 CRC64;

Query Match 100.0%; Score 372; DB 4; Length 448;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVLSPLSFIRKSHVMOGILLVFAKYQHLPYIQLSTKSTPGTGLFGYGNKGGVNICL 60  
|||||  
77 MDVLSPLSFIRKSHVMOGILLVFAKYQHLPYIQLSTKSTPGTGLFGYGNKGGVNICL 136  
61 KLYGYYSIINCHLPKPHISNNYQRLHFDRILEMONCEGRDIPNLDHLLIIFGDMNFR 120  
137 KLYGYYSIINCHLPKPHISNNYQRLHFDRILEMONCEGRDIPNLDHLLIIFGDMNFR 196  
QY 121 IEDGGLHFRRESIKNRCYGLMEKDQLSIKKHDPLEFQGRLLFPPTYKFDNSNDY 180  
|||||  
197 IEDGGLHFRRESIKNRCYGLMEKDQLSIKKHDPLEFQGRLLFPPTYKFDNSNDY 256  
QY 181 DTSEKRRKPAWTDRLMLRKQPCAGPDTPIPPASHFSLSLRGYSSHMTYGISDHPKPVSG 240  
|||||  
257 DTSEKRRKPAWTDRLMLRKQPCAGPDTPIPPASHFSLSLRGYSSHMTYGISDHPKPVSG 316  
QY 241 TFDELEKPLVASPLVLMPEDLMTVENDMVSYSTSDPSSPMDWIGLYKVGGLRDVNDY 300  
|||||  
317 TFDELEKPLVASPLVLMPEDLMTVENDMVSYSTSDPSSPMDWIGLYKVGGLRDVNDY 376  
QY 301 VSYAMVGDGSKVSCSDNLNQYVIDISNIPPTTEDEFLLCYSSLSRVGVSIRPQIPPSL 360  
|||||  
377 VSYAMVGDGSKVSCSDNLNQYVIDISNIPPTTEDEFLLCYSSLSRVGVSIRPQIPPSL 436  
QY 361 REDPLGEAOPQI 372  
|||||  
DB 437 REDPLGEAOPQI 448

RESULT 2  
ID 09P2R5 PRELIMINARY; PRT: 372 AA.  
09P2R5: 01-OCT-2000 (TREMBLrel. 15, Created)  
01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE 43-KDA FORM SKELETAL MUSCLE AND KIDNEY ENRICHED INOSITOL  
PHOSPHATASE.  
GN SKIP 43-KDA FORM.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
NC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ijuin T., Mochizuki Y., Fukami K., Funaki M., Asano T., Takenawa T.;  
RT "Identification and Characterization of a Novel Inositol Polyphosphate  
5-Phosphatase."  
RL J. Biol. Chem. 0:0-0(2000).  
DR EMBL: AB036830; BAA92341.1; -  
DR InterPro: IPR000300; IPPC:  
DR Pfam: PF00783; IPPC: 1.  
DR SMART: SM00128; IPPC: 1.  
SQ SEQUENCE 372 AA; 42922 MW; A2FCAE390DCD4528 CRC64;

Query Match 91.1%; Score 339; DB 4; Length 372;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVLSPLSFIRKSHVMOGILLVFAKYQHLPYIQLSTKSTPGTGLFGYGNKGGVNICL 60  
|||||  
Db 1 MDVLSPLSFIRKSHVMOGILLVFAKYQHLPYIQLSTKSTPGTGLFGYGNKGGVNICL 60  
QY 61 KLYGYYSIINCHLPKPHISNNYQRLHFDRILEMONCEGRDIPNLDHLLIIFGDMNFR 120  
|||||  
Db 61 KLYGYYSIINCHLPKPHISNNYQRLHFDRILEMONCEGRDIPNLDHLLIIFGDMNFR 120  
QY 121 IEDGGLHFRRESIKNRCYGLMEKDQLSIKKHDPLEFQGRLLFPPTYKFDNSNDY 180  
|||||  
Db 121 IEDGGLHFRRESIKNRCYGLMEKDQLSIKKHDPLEFQGRLLFPPTYKFDNSNDY 180  
QY 181 DTSEKRRKPAWTDRLMLRKQPCAGPDTPIPPASHFSLSLRGYSSHMTYGISDHPKPVSG 240  
|||||  
Db 181 DTSEKRRKPAWTDRLMLRKQPCAGPDTPIPPASHFSLSLRGYSSHMTYGISDHPKPVSG 240  
QY 241 TFDELEKPLVASPLVLMPEDLMTVENDMVSYSTSDPSSPMDWIGLYKVGGLRDVNDY 300  
|||||  
Db 241 TFDELEKPLVASPLVLMPEDLMTVENDMVSYSTSDPSSPMDWIGLYKVGGLRDVNDY 300  
QY 301 VSYAMVGDGSKVSCSDNLNQYVIDISNIPPTTEDEFLLCY 339  
|||||  
Db 301 VSYAMVGDGSKVSCSDNLNQYVIDISNIPPTTEDEFLLCY 339

RESULT 3  
ID 09NPJ5 PRELIMINARY; PRT: 448 AA.  
09NPJ5: 01-OCT-2000 (TREMBLrel. 15, Created)  
01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE SKELETAL MUSCLE AND KIDNEY ENRICHED INOSITOL PHOSPHATASE.  
GN SKIP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
NC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ijuin T., Mochizuki Y., Fukami K., Funaki M., Asano T., Takenawa T.;

\*Identification and Characterization of a Novel Inositol Polyphosphate

RT 5-Phosphatase.  
RT J. Biol. Chem. 0:0-0(2000).  
DR EMBL: AB036831; BAA92342.1; -  
DR EMBL: AB036829; BAA92340.1; -  
DR InterPro: IPR000300; IPIC.  
DR Pfam: PF00783; IPIC; 1.  
DR SMART: SM00128; IPIC; 1.  
SQ SEQUENCE 448 AA; 51228 MW; 0651BE3CE2EEB0F CRC64;

Query Match 91.1%; Score 339; DB 4; Length 448;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVLSPLSFIVKSHVRKQGIILLVFAKYOHLPYIOILSTKSTPTGLGVMGNKGVNCL 60  
DB 77 MDVLSPLSFIVKSHVRKQGIILLVFAKYOHLPYIOILSTKSTPTGLGVMGNKGVNCL 136  
QY 61 KLYGYVSIINCHLPPHISNNYQRLHEHFDRIEMQNEGRDIPNILDHLLIWFEGDNFR 120  
DB 137 KLYGYVSIINCHLPPHISNNYQRLHEHFDRIEMQNEGRDIPNILDHLLIWFEGDNFR 196  
QY 121 IEQGLHFVRESIKKRCYGGIMKQDQSIKHPDLREFQEGFLPPYTKFPRNSNDY 180  
DB 197 IEQGLHFVRESIKKRCYGGIMKQDQSIKHPDLREFQEGFLPPYTKFPRNSNDY 256  
QY 181 DTSEKKRPAMTDRIILRLKQPCAGDPTPIPPASHFSLRGSSHMTYGISPHKPYSG 240  
DB 257 DTSEKKRPAMTDRIILRLKQPCAGDPTPIPPASHFSLRGSSHMTYGISPHKPYSG 316  
QY 241 TFDLELPLVSAPLIVLMPEDLMTVENDMVSYSTSDPSSPMDWIGLYKVLGVNDY 300  
DB 317 TFDLELPLVSAPLIVLMPEDLMTVENDMVSYSTSDPSSPMDWIGLYKVLGVNDY 376  
QY 301 VSYAWGDSKVSQSDNLNOYIDISNPTDEFLCY 339  
DB 377 VSYAWGDSKVSQSDNLNOYIDISNPTDEFLCY 415

RESULT 4  
Q15733 PRELIMINARY; PRT; 329 AA.  
AC Q15733;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE PHOSPHATIDYLINOSITOL (4,5)BISPHOSPHATE 5-PHOSPHATASE HOMOLOG (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RA Nussbaum R.L.;  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U45973; AAB03214.1; -  
DR InterPro: IPR000300; IPIC.  
DR Pfam: PF00783; IPIC; 1.  
DR SMART: SM00128; IPIC; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 329 AA; 37996 MW; 834B095B03BB06D5 CRC64;

Query Match 79.3%; Score 295; DB 4; Length 329;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLFGYMGNGGVNCLKLYGYVSIINCHLPPHISNNYQRLHEHFDRIEMQNEGRDIPN 104  
DB 2 GLFGYMGNGGVNCLKLYGYVSIINCHLPPHISNNYQRLHEHFDRIEMQNEGRDIPN 61

QY 105 ILDHLLIWFEGDNFRIEQGLHFVRESIKKRCYGGIMKQDQSIKHPDLREFQEGR 164  
DB 62 ILDHLLIWFEGDNFRIEQGLHFVRESIKKRCYGGIMKQDQSIKHPDLREFQEGR 121  
QY 165 LFPPTKFPDRNSNDYDTSEKKRPAMTDRIILRLKQPCAGDPTPIPPASHFSLRGY 224  
DB 122 LFPPTKFPDRNSNDYDTSEKKRPAMTDRIILRLKQPCAGDPTPIPPASHFSLRGY 181  
QY 225 SSHMTYGISDHPKPVSGTFDELKPLVSAPLIVLMPEDLMTVENDMVSYSTSDPSSPW 284  
DB 182 SSHMTYGISDHPKPVSGTFDELKPLVSAPLIVLMPEDLMTVENDMVSYSTSDPSSPW 241  
QY 285 DWIGLYKVLGVNDYVSYAWGDSKVSQSDNLNOYIDISNPTDEFLCY 339  
DB 242 DWIGLYKVLGVNDYVSYAWGDSKVSQSDNLNOYIDISNPTDEFLCY 296

RESULT 5  
Q09040 PRELIMINARY; PRT; 468 AA.  
AC Q09040;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE PUTATIVE PHOSPHATASE (PUTATIVE PHOSPHOINOSITIDE 5-PHOSPHATASE TYPE II).  
GN PPS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DBA/2J;  
RA Hamilton B.A., Smith D.J., Mueller K.L., Kerrebrock A.W.,  
RA Bronson R.T., van Berkel V., Daly M.J., Kruglyak L., Reeve M.P.,  
RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
RL EMBL: U96726; AAC60757.1; -  
DR EMBL: U96726; AAC60757.1; -  
DR MGI: 1194899; PPS.  
DR InterPro: IPR000300; IPIC.  
DR Pfam: PF00783; IPIC; 1.  
DR SMART: SM00128; IPIC; 1.  
SQ SEQUENCE 468 AA; 54158 MW; F2E1CA370B97A8A1 CRC64;

Query Match 6.5%; Score 24; DB 11; Length 468;  
Best Local Similarity 100.0%; Pred. No. 1,6e-16;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 YDTSEKKRPAMTDRIILRLKQPCAGDPTPIPPASHFSLRGY 203  
DB 274 YDTSEKKRPAMTDRIILRLKQPCAGDPTPIPPASHFSLRGY 297

RESULT 6  
Q15735 PRELIMINARY; PRT; 397 AA.  
AC Q15735;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE PHOSPHATIDYLINOSITOL (4,5)BISPHOSPHATE 5-PHOSPHATASE HOMOLOG (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;

RA Nussbaum R.L.;  
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U45975; AAB03216.1; -  
 DR InterPro: IPR000300; IPPC.  
 DR Pfam: PF00783; IPPC; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 397 AA; 43893 MW; 71418E311E24FBFA CRC64;

Query Match 3.5%; Score 13; DB 4; Length 397;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 185 KKRKPAMTDRIW 197  
 Db 63 KKRKPAMTDRIW 75

RESULT 7

QJUMC1 PRELIMINARY; PRT; 1001 AA.

DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE POLYLINE-RICH INOSITOL POLYPHOSPHATE 5-PHOSPHATASE.  
 GN PIP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE=20062911; PubMed=10593988;  
 RA Mochizuki Y.; Takenawa T.;  
 RT "Novel inositol polyphosphate 5-phosphatase localizes at membrane  
 ruffles.";  
 RL J. Biol. Chem. 274:36790-36795(1999).  
 DR EMBL; AB032551; BAA90553.1; -  
 DR InterPro: IPR000300; IPPC.  
 DR InterPro: IPR002965; P\_rich\_extensn.  
 DR Pfam: PF00783; IPPC; 1.  
 DR PRINTS; PR01217; PRICHEXTENS.  
 DR SMART; SM00128; IPPC; 1.  
 SQ SEQUENCE 1001 AA; 107207 MW; 7BE741FEF8E3FAB CRC64;

Query Match 3.5%; Score 13; DB 11; Length 1001;  
 Best Local Similarity 100.0%; Pred. No. 0.00013;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 185 KKRKPAMTDRIW 197  
 Db 667 KKRKPAMTDRIW 679

RESULT 8

Q9UDT9 PRELIMINARY; PRT; 1056 AA.

DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE WIGSC:H DJ412A9.2 PROTEIN (FRAGMENT).  
 GN WIGSC:H DJ412A9.2  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 DR MEDLINE=99063792; PubMed=9847074;

RA Sulston J.E., Waterston R.;  
 RL "Toward a complete human genome sequence.";  
 RT Genome Res. 8:1097-1108(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Murray J., Lennox S., Harmon G.;  
 RT "The sequence of Homo sapiens PAC clone RP3-412A9.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.

RA Waterston R.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC005005; AAD15618.1; -  
 DR InterPro: IPR000300; IPPC.  
 DR InterPro: IPR002965; P\_rich\_extensn.  
 DR Pfam: PF00783; IPPC; 2.  
 DR PRINTS; PR01217; PRICHEXTENS.  
 DR SMART; SM00128; IPPC; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 1056 AA; 112564 MW; B82A8C134D72F7D0 CRC64;

Query Match 3.5%; Score 13; DB 4; Length 1056;  
 Best Local Similarity 100.0%; Pred. No. 0.00014;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 185 KKRKPAMTDRIW 197  
 Db 722 KKRKPAMTDRIW 734

RESULT 9

Q9XUD3 PRELIMINARY; PRT; 1113 AA.

ID Q9XUD3  
 AC Q9XUD3; Q9GF41;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Jc8.10A PROTEIN (SYNAPPOANIN UNC-26B).  
 GN Jc8.10A OR UNC-26.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lightning J.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C. elegans: A platform for  
 investigating biology.";  
 RL Science 282:2012-2018(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20391980; PubMed=10931870;  
 RA Harris T.W., Hartweg E., Horvitz H.R., Jorgensen E.M.;  
 RT "Mutations in synaptojanin disrupt synaptic vesicle recycling.";  
 RL J. Cell Biol. 150:589-600(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Sulston J.E.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z82274; CAB05234.2; -  
 DR EMBL; AL132951; CAB05234.2; JOINED.  
 DR EMBL; AF283323; AAG18575.1; -  
 DR EMBL; AL132951; CAC44311.1; -  
 DR EMBL; Z82274; CAC44311.1; JOINED.  
 DR InterPro: IPR000300; IPPC.  
 DR InterPro: IPR002013; SYJA\_N.  
 DR Pfam: PF00783; IPPC; 1.

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 09:22:59 ; Search time 31.53 Seconds

(without alignments)  
1310.481 Million cell updates/sec

Title: US-09-892-287-1

Perfect score: 372

Sequence: 1 MDVLSPLSTIKSHVMQCI.....FQIPPSLRDPDPEAQPPI 372

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

W size: 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : A\_Geneseq\_032802.\*

1: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	372	100.0	372	20	AAW97094
2	13	3.5	255	21	AAW27797
3	13	3.5	279	22	AAW33981
4	13	3.5	381	21	AAW27845
5	13	3.5	382	21	AAW27846
6	8	2.2	95	20	AAV35802
7	7	1.9	68	22	AAW88231
8	7	1.9	74	22	AAW44115
9	7	1.9	76	22	AAW20728
10	7	1.9	95	22	AAW01673
11	7	1.9	108	22	AAW30209

12	7	1.9	121	21	AAW51521	Arabidopsis thalia
13	7	1.9	148	21	AAW57447	Arabidopsis thalia
14	7	1.9	148	21	AAW59665	Arabidopsis thalia
15	7	1.9	151	21	AAW57446	Arabidopsis thalia
16	7	1.9	151	21	AAW59664	Arabidopsis thalia
17	7	1.9	152	21	AAW51520	Arabidopsis thalia
18	7	1.9	172	21	AAW51521	Arabidopsis thalia
19	7	1.9	173	21	AAW74640	Arabidopsis thalia
20	7	1.9	176	22	AAW16962	Human novel secret
21	7	1.9	178	22	AAW42117	Human novel secret
22	7	1.9	179	21	AAW57445	Arabidopsis thalia
23	7	1.9	179	21	AAW59663	Arabidopsis thalia
24	7	1.9	192	22	AAW53242	Human breast cancer
25	7	1.9	193	21	AAW74639	Neisseria meningit
26	7	1.9	193	21	AAW74641	Neisseria meningit
27	7	1.9	202	21	AAW51519	Arabidopsis thalia
28	7	1.9	207	8	AAW70588	DNA sequence encod
29	7	1.9	207	8	AAW70589	DNA sequence encod
30	7	1.9	207	11	AAW05495	Thaumatin 1, Thau
31	7	1.9	207	17	AAW85303	Thaumatin 1, Thau
32	7	1.9	208	6	AAW50285	Sequence of a poly
33	7	1.9	223	15	AAW54641	HCV EN-80-2 nucleo
34	7	1.9	223	18	AAW07484	HCV unprocessed co
35	7	1.9	223	18	AAW18199	Hepatitis C virus
36	7	1.9	223	18	ABG14027	Novel human diagno
37	7	1.9	224	14	AAW43068	HCV antigen, Hepa
38	7	1.9	224	15	AAW8277	Hepatitis C virus
39	7	1.9	229	22	AAW17012	Human novel secret
40	7	1.9	229	22	AAW75098	Human colon cancer
41	7	1.9	235	3	AAW20003	Preprothymalatin a
42	7	1.9	258	20	AAW41212	E. coli MtrB polyp
43	7	1.9	258	21	AAW93223	An Escherichia coli
44	7	1.9	335	21	AAW40331	Human polypeptide
45	7	1.9	351	21	AAW1589	Streptococcus pneu
46	7	1.9	357	21	AAW1589	Human ORFX ORF149
47	7	1.9	370	21	AAW1655	Human ORFX ORF149
48	7	1.9	388	21	AAW22240	Arabidopsis thalia
49	7	1.9	388	21	AAW54446	Arabidopsis thalia
50	7	1.9	395	21	AAW32239	Arabidopsis thalia
51	7	1.9	395	21	AAW5445	Arabidopsis thalia
52	7	1.9	396	22	AAW17034	Human novel secret
53	7	1.9	425	21	AAW22238	Arabidopsis thalia
54	7	1.9	425	21	AAW54444	Arabidopsis thalia
55	7	1.9	428	22	ABW1914	Drosophila melanog
56	7	1.9	470	14	AAW34481	Encoded by Hepatit
57	7	1.9	484	22	AAW36429	Pseudomonas aerugi
58	7	1.9	508	22	ABW4662	Drosophila melanog
59	7	1.9	578	22	AAW5820	Haemophilus influe
60	7	1.9	643	22	ABW6087	Drosophila melanog
61	7	1.9	649	22	ABW5681	Novel human diagno
62	7	1.9	749	22	ABW5181	Human protein sequ
63	7	1.9	1218	22	ABW62412	Drosophila melanog
64	7	1.9	1218	22	ABW6736	Drosophila melanog
65	7	1.6	8	14	AAW35978	Hepatitis C virus
66	6	1.6	8	14	AAW35979	Hepatitis C virus
67	6	1.6	8	14	AAW35979	Hepatitis C virus
68	6	1.6	9	13	AAW26496	Fragment of open r
69	6	1.6	9	15	AAW47636	Hepatitis C virus
70	6	1.6	9	16	AAW6718	NS4-region epitope
71	6	1.6	9	18	AAW39722	Human carcino-emb
72	6	1.6	9	19	AAW70050	CEA derived HLA-A2
73	6	1.6	9	20	AAW46534	Immunogenic peptid
74	6	1.6	9	20	AAW47656	Immunogenic peptid
75	6	1.6	9	20	AAW6177	IKK-alpha polypept
76	6	1.6	9	21	AAW51432	Hepatitis C virus
77	6	1.6	10	17	AAW00681	Peptide comprising
78	6	1.6	12	14	AAW80657	Sequence of peptid
79	6	1.6	12	17	AAW08273	HCV NS4 region epi
80	6	1.6	12	17	AAW03058	Epitope from HCV N
81	6	1.6	12	17	AAW88795	HCV NS4 epitope 1a
82	6	1.6	14	17	AAW03352	HCV NS4/1 epitope
83	6	1.6	14	21	AAW82134	Monoclonal antibod
84	6	1.6	14	22	AAW88123	CD66 peptide CD66

85 6 1.6 15 21 AAY98930 HLA class II bindi  
86 6 1.6 16 15 AAR51409 Branched peptide C  
87 6 1.6 17 14 AAR37946 HCV NS-4 type 1 re  
88 6 1.6 17 14 AAR37947 HCV NS-4 type 1 re  
89 6 1.6 20 18 AAM10133 Hepatitis C virus  
90 6 1.6 20 18 AAM10134 Hepatitis C virus  
91 6 1.6 20 19 AAM47151 Hepatitis C virus  
92 6 1.6 20 21 AAY57210 NS4A/4B cleavage s  
93 6 1.6 21 20 AAY12754 Human 5' EST seque  
94 6 1.6 22 14 AAR41108 HCV peptide HCV3.  
95 6 1.6 22 14 AAR41109 HCV peptide X or H  
96 6 1.6 22 14 AAR41110 HCV peptide XI or  
97 6 1.6 23 16 AAR70220 Hepatitis C virus  
98 6 1.6 23 16 AAR70221 Hepatitis C virus  
99 6 1.6 25 15 AAR51403 Branched peptide C  
100 6 1.6 26 14 AAR41183 HCV NS4 protein HC

## ALIGNMENTS

## RESULT 1

ID AAM97094 standard; Protein: 372 AA.

AC AAM97094:

DT 28-APR-1999 (first entry)

DE Phosphatidylinositol 4, 5-bisphosphate 5-phosphatase.

KW Human; phosphatidylinositol 4,5-bisphosphate 5-phosphatase; PBPP;

KW Incyte clone 638789; antagonist; immune disorder; cancer;

KW neuronal disorder; human tubby homologue.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Modified-site 38 /note= "potential phosphorylation site"  
FT Modified-site 132 /note= "potential phosphorylation site"  
FT Modified-site 170 /note= "potential phosphorylation site"  
FT Modified-site 183 /note= "potential phosphorylation site"  
FT Modified-site 192 /note= "potential phosphorylation site"  
FT Modified-site 275 /note= "potential phosphorylation site"  
FT Modified-site 282 /note= "potential phosphorylation site"  
FT Modified-site 295 /note= "potential phosphorylation site"  
FT Modified-site 312 /note= "potential phosphorylation site"  
FT Modified-site 329 /note= "potential phosphorylation site"  
FT Modified-site 330 /note= "potential phosphorylation site"  
FT Modified-site 359 /note= "potential phosphorylation site"  
FT Modified-site /note= "potential phosphorylation site"

PN WO9900507-A1.

PD 07-JAN-1999.

PF 26-JUN-1998; 98WO-US13399.

PR 27-JUN-1997; 97US-0884681.

PA (INCYTE) INCYTE PHARM INC.

PI Corley NC, Hillman JL, Lal P, Shah P:  
XX WPI: 1999-095752/08.  
DR N-PSDB; AAX15254.

PT Phosphatidylinositol 4,5-bisphosphate 5-phosphatase - used for  
FT treating immune disorders, cancers, and neuronal disorders  
PS Claim 1; Fig 1A-G; 81pp; English.

XX The present sequence encodes a human phosphatidylinositol  
CC 4,5-bisphosphate 5-phosphatase (PBPP) protein. The PBPP polynucleotide  
CC was first identified in Incyte clone 638789 from the breast cancer  
CC cDNA library BRSTN03. Antagonists of the PBPP protein can be used  
CC in the treatment or prevention of an immune disorder, a cancer, or  
CC a neuronal disorder. The PBPP polynucleotide can be used for the  
CC detection of polynucleotides encoding human tubby homologue. The immune  
CC disorders that can be treated include AIDS, Addison's disease,  
CC adult respiratory distress syndrome, allergies, anaemia, asthma,  
CC atherosclerosis, Crohn's disease, ulcerative colitis, atopic dermatitis,  
CC gout, Grave's disease, irritable bowel syndrome, lupus erythematosus,  
CC multiple sclerosis, myasthenia gravis, osteoarthritis, osteoporosis,  
CC rheumatoid arthritis, scleroderma, and autoimmune thyroiditis. Also,  
CC complications of cancer, haemodialysis, extracorporeal circulation,  
CC infection, and trauma can be treated using the antagonist. The neuronal  
CC disorders that can be treated include Alzheimer's disease, amnesia,  
CC cataplexia, amyotrophic lateral sclerosis, dementia, depression,  
CC Down's syndrome, epilepsy, Huntington's disease, multiple sclerosis,  
CC neurofibromatosis, Parkinson's disease, paranoid psychoses, schizophrenia  
CC and Tourette's syndrome.

SQ Sequence 372 AA:

Query Match 100.0%; Score 372; DB 20; Length 372;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVLSPLSFIRKVSHPVMOGILLVFAKYOHLPYIOILSTKSTPGGLFGYMGNGKGVNLC 60  
DB 1 mdvlsplsfirkvs hvmrgmlllvfakypohlp yioilstkstpgglfg ymgngkgvnlc 60  
QY 61 KLYGYVSIINCHLPPIHSNNYQRIEHPDRILEMONCGRQIPNLDDLIITWGDMMFR 120  
DB 61 klygyvsiinchlp pihsnnyqriehpdrile moncgrqipnl ddiitwgdmmfr 120  
QY 121 IEDFGLHFVRESIKRRCYGLMEKQDLSIAKHPDLREFQEGRLFPPTYKFRNSNDY 180  
DB 121 iedfghlfvresik rrcyglmekqdlsiak hpdrlrefqegrlf ptykfrnsndy 180  
QY 181 DTSEKRRKPAWTDRIILWLRKROPAGPDPPIPPASHFSLRGYSSHMTYGISDHKPVSG 240  
DB 181 dtsekkrrkpa wtdriilwlrkropagpdp pippashfslr gysshmtygisdhkp vsg 240  
QY 241 TFDELEKPLVSAPLIVLMPEDLMTVENDMVSYSTSPSSPMDMIGLYRGLRDVNDY 300  
DB 241 tfdelekplvsa plivlmpedlmtven dmvsystspsspmdm iglyrglrdvndy 300  
QY 301 VSYAMVGSKVCSDNLNOVYIDISNIPTDEFLACYNSLSVSGISPPFOIPCSL 360  
DB 301 vsyamvgskvc sdnlnovyidisnipt deflacyns lsvsgisppfo ipcs 360  
QY 361 REDPLGEAOPQI 372  
DB 361 redplgeaopqi 372

## RESULT 2

ID AAB27797 standard; Protein: 255 AA.

AC AAB27797:



DT 29-JAN-2001 (first entry)  
XX Human secreted protein #4.  
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.  
XX Homo sapiens.  
OS  
XX  
PN WO200055199-A1.  
XX  
PD 21-SEP-2000.  
XX  
PF 09-MAR-2000; 2000WO-US06014.  
XX  
PR 12-MAR-1999; 99US-0124095.  
PR 11-JUN-1999; 99US-0138598.  
PR 03-DEC-1999; 99US-0168665.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM, Komatsoulis G;  
XX  
DR WPI: 2000-572359/53.  
DR N-PSDB; AAC59218.  
XX  
PT Isolated nucleic acid molecule encoding a human secreted protein is  
XX used in preventing, treating or ameliorating a medical condition -  
XX  
PS Claim 11; Page 372-373; 433pp; English.  
XX  
CC Sequences AAB27794-B27840 represent the amino acid sequences of 47  
CC human secreted proteins encoded by the genes AAC59215-59261. The genes  
CC and proteins are useful for preventing, ameliorating or treating medical  
CC conditions, e.g. by protein or gene therapy. The genes are isolated from  
CC a range of human tissues disclosed in the specification. The nucleic  
CC acids, proteins, antibodies and (ant)agonists are useful in the  
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,  
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
XX  
SQ Sequence 255 AA;  
XX  
Query Match 3.5%; Score 13; DB 21; Length 255;  
Best Local Similarity 100.0%; Pred. No. 0.0004;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 185 KRRKPAMTDRIILW 197  
DB 169 krrkpawtdrllw 181  
XX  
RESULT 3  
AAG73981 ID AAG73981 standard; Protein; 279 AA.  
XX  
AC AAG73981;  
XX  
DT 03-SEP-2001 (first entry)  
XX  
DE Human colon cancer antigen protein SEQ ID NO:4745.  
XX

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma.  
XX  
XX  
OS Homo sapiens.  
XX  
PN WO200122920-A2.  
XX  
PD 05-APR-2001.  
XX  
PF 28-SEP-2000; 2000WO-US26524.  
XX  
PR 29-SEP-1999; 99US-0157137.  
PR 03-NOV-1999; 99US-0163280.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX  
DR WPI: 2001-235357/24.  
DR N-PSDB; AAH33412.  
XX  
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
XX useful for preventing, diagnosing and/or treating colorectal cancers -  
XX  
PS Claim 11; Page 6545-6546; 9803pp; English.  
XX  
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patients own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAB77789 represent sequences used in the exemplification of the  
CC present invention.  
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
XX  
SQ Sequence 279 AA;  
XX  
Query Match 3.58%; Score 13; DB 22; Length 279;  
Best Local Similarity 100.0%; Pred. No. 0.00044;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 185 KRRKPAMTDRIILW 197  
DB 194 krrkpawtdrllw 206  
XX  
RESULT 4  
AAB27845 ID AAB27845 standard; Protein; 381 AA.  
XX  
AC AAB27845;  
XX  
DT 29-JAN-2001 (first entry)  
XX  
DE Sequence homologous to protein fragment encoded by gene 4.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.

XX OS Homo sapiens.  
XX PN WO20005199-A1.  
XX PD 21-SEP-2000.  
XX PF 09-MAR-2000; 2000WO-US06014.  
XX PR 12-MAR-1999; 99US-0124095.  
XX PR 11-JUN-1999; 99US-0138598.  
XX PR 03-DEC-1999; 99US-0168665.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Ruben SM, Komatsoulis G;  
XX DR WPI: 2000-572359/53.  
XX  
PS Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition -  
PS Disclosure; Page 491-493; 433pp; English.  
XX CC The invention relates to the isolation of genes AAC59215-C59261 encoding the human secreted proteins AAB27794-B27840. This sequence represents a peptide fragment homologous to the protein encoded by the gene given in the descriptor line. The sequence is a search result from a BLASTX homology search. The genes and proteins are useful for preventing, CC ameliorating or treating medical conditions, e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and CC (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, CC lung, or urogenital; (b) immune disorders e.g. Addison's disease, CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, CC diabetes mellitus, Crohn's disease, multiple sclerosis, Rheumatoid CC arthritis and ulcerative colitis; (c) cardiovascular disorders such CC as myocardial ischaemias; (d) wound healing; (e) neurological diseases CC e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such CC as viral, bacterial, fungal and parasitic infections.  
XX SQ Sequence 381 AA;  
SQ  
Query Match 3.5%; Score 13; DB 21; Length 381;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 185 KRRPAMTDRIILW 197  
|||  
Db 60 KRRPAMTDRIILW 72  
RESULT 5  
AAB27846  
ID AAB27846 standard; Protein: 382 AA.  
XX AC AAB27846;  
XX DT 29-JAN-2001 (first entry)  
XX DE Protein fragment encoded by gene 4.  
XX DE Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.  
XX OS Homo sapiens.  
XX

PN WO20005199-A1.  
XX PD 21-SEP-2000.  
XX PF 09-MAR-2000; 2000WO-US06014.  
XX PR 12-MAR-1999; 99US-0124095.  
XX PR 11-JUN-1999; 99US-0138598.  
XX PR 03-DEC-1999; 99US-0168665.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Ruben SM, Komatsoulis G;  
XX DR WPI: 2000-572359/53.  
XX  
PS Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition -  
PS Disclosure; Page 403-404; 433pp; English.  
XX CC The invention relates to the isolation of genes AAC59215-C59261 encoding the human secreted proteins AAB27794-B27840. This sequence represents a fragment of the protein encoded by the gene given in the descriptor line. The sequence is used as a query sequence for doing BLASTX searches to determine homologous sequence to the protein. The genes and proteins are useful for preventing, ameliorating or treating medical conditions, e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, CC proteins, antibodies and (ant)agonists are useful in the diagnosis, CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer, CC and other cancers of the adrenal gland, bone, bone marrow, breast, CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c) CC cardiovascular disorders such as myocardial ischaemias; (d) wound CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and CC (f) infectious diseases such as viral, bacterial, fungal and parasitic CC infections.  
XX SQ Sequence 382 AA;  
SQ  
Query Match 3.5%; Score 13; DB 21; Length 382;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 185 KRRPAMTDRIILW 197  
|||  
Db 60 KRRPAMTDRIILW 72  
RESULT 6  
AAV35802  
ID AAV35802 standard; Protein: 95 AA.  
XX AC AAV35802;  
XX DT 13-SEP-1999 (first entry)  
XX DE Chlamydia pneumoniae protein not found in C. trachomatis.  
XX DE Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
KW vaccine; neutralising epitope.  
XX Chlamydia pneumoniae.  
XX WO9927105-A2.  
XX PN 03-JUN-1999.  
XX PD

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 09:19:23 ; Search time 18.61 Seconds  
(without alignments)  
488.249 Million cell updates/sec

Title: US-09-892-287-1  
Perfect score: 2020  
Sequence: 1 MDVLSPLSFVKVSHVRMGL.....FOIPGSLREDPLGPAQPI 372

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 231628 seqs, 2442594 residues

Number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	%			Length	DB	ID	Description		
	Score	Match	Query						
1	2020	100.0	372	2	US-08-884-681-1		Sequence 1, Appli		
2	2020	100.0	372	4	US-09-258-643-1		Sequence 1, Appli		
3	1791	88.7	329	2	US-08-884-681-3		Sequence 3, Appli		
4	1791	88.7	329	4	US-09-258-643-3		Sequence 3, Appli		
5	432.5	21.4	942	2	US-08-884-681-4		Sequence 4, Appli		
6	432.5	21.4	942	3	US-08-560-005-6		Sequence 6, Appli		
7	432.5	21.4	942	4	US-09-258-643-4		Sequence 4, Appli		
8	432.5	21.4	942	4	US-09-418-540-6		Sequence 6, Appli		
9	368.5	18.2	968	3	US-08-560-005-7		Sequence 7, Appli		
10	368.5	18.2	968	4	US-09-418-540-7		Sequence 7, Appli		
11	367.5	18.2	901	2	US-08-884-681-5		Sequence 5, Appli		
12	367.5	18.2	901	4	US-09-258-643-5		Sequence 5, Appli		
13	290.5	14.4	946	3	US-08-560-005-4		Sequence 4, Appli		
14	290.5	14.4	946	4	US-09-418-540-4		Sequence 4, Appli		
15	287	14.2	1149	3	US-08-560-005-5		Sequence 5, Appli		
16	287	14.2	1149	4	US-09-418-540-5		Sequence 5, Appli		
17	263	13.0	654	3	US-08-560-005-10		Sequence 10, Appl		
18	263	13.0	654	4	US-09-418-540-10		Sequence 10, Appl		
19	252	12.5	968	1	US-08-434-730-14		Sequence 14, Appl		
20	251	12.4	1185	4	US-08-664-962B-2		Sequence 2, Appli		
21	251	12.4	1185	4	US-09-311-743-2		Sequence 2, Appli		
22	250	12.4	976	3	US-08-560-005-2		Sequence 2, Appli		
23	250	12.4	976	3	US-09-195-868-14		Sequence 14, Appl		
24	250	12.4	976	4	US-09-418-540-2		Sequence 2, Appli		
25	250	12.4	1187	4	US-08-664-962B-8		Sequence 8, Appli		
26	250	12.4	1187	4	US-09-311-743-8		Sequence 8, Appli		
27	250	12.4	1189	3	US-09-195-868-15		Sequence 15, Appl		

28 250 12.4 1229 3 US-09-195-868-28 Sequence 28, Appli  
29 220.5 10.9 398 3 US-08-560-005-3 Sequence 3, Appli  
30 220.5 10.9 398 4 US-09-418-540-3 Sequence 3, Appli  
31 110 5.4 121 3 US-08-560-005-8 Sequence 8, Appli  
32 110 5.4 121 4 US-09-418-540-8 Sequence 8, Appli  
33 98.5 4.9 412 3 US-08-560-005-9 Sequence 9, Appli  
34 98.5 4.9 412 4 US-09-418-540-9 Sequence 9, Appli  
35 87 4.3 735 1 US-08-021-601-4 Sequence 4, Appli  
36 87 4.3 735 1 US-08-082-849B-4 Sequence 4, Appli  
37 87 4.3 735 5 PCT-US94-01624-4 Sequence 4, Appli  
38 87 4.3 903 1 US-08-021-601-12 Sequence 12, Appli  
39 87 4.3 903 1 US-08-082-849B-12 Sequence 12, Appli  
40 87 4.3 903 5 PCT-US94-01624-12 Sequence 12, Appli  
41 83 4.1 477 3 US-08-704-711A-20 Sequence 20, Appli  
42 83 4.1 477 4 US-08-281-313-1 Sequence 9, Appli  
43 82.5 4.1 239 1 US-08-103-998-2 Sequence 2, Appli  
44 82.5 4.1 622 2 US-08-664-646A-2 Sequence 2, Appli  
45 82.5 4.1 622 2 US-09-066-285-2 Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-08-884-681-1  
; Sequence 1, Application US/08884681  
; Patent No. 5953338  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: NEW HUMAN PHOSPHATIDYLINOSITOL  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/884,681  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0334 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 372 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BRSTNOT03  
; CLONE: 638789  
; US-08-884-681-1

Query Match 100.0%; Score 2020; DB 2; Length 372;  
Best Local Similarity 100.0%; Pred. No. 3.5e-210;







```

: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/884,681
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0334 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 942 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 1019103
: US-09-258-643-4

```

Query Match 21.4% Score 432.5; DB 4; Length 942;  
Best Local Similarity 37.0% Pred. No. 1.1e-37;  
Matches 90; Conservative 45; Mismatches 87; Gaps 4;  
Indels 21;

Qy	9	FIKVSHVRMOGILLLLVFAKQVHLPYTOILSTKSTPTGLFGYMGKGGVNICLKLYGYYS	68
Db	365	YAKVKILRLVIGIMLLLVKQEAHYSEVEAEFTVGTGIMGRMGKGVAREQFHNSTIC	424
Qy	69	IINCHLPPHSTNNYQRLHF--DRILEMONE--GRDIP--NILDHLLIWFQDWNRIEDF	124
Db	425	VVNSHLAAHIEEYERNQYKDCSRMQFCQPPSPPLTTISNHDVILWGLDUNYRIEL	484
Qy	125	GLHFVRESTKNRCYGLWEKQDLSIAKKHDPDLLREFQEGRLFPPTYKFDNSNDYDTE	184
Db	485	DVEKVKVLLIEEKDFQMLAYDQLKIQVAARTVPEGETEGELTFQPTYKYDTGSDWDTSE	544
Qy	185	KKKPAWTDRLRLWRKQPCAGPDTTIPASHFSLRQYSSHMTYGISDHKPVSGTDFL	244
Db	545	KCRAPAWCDRLWKVK-----NITQLSYQSHMAUKTSDHKRPVSSVFDI	587
Qy	245	ELK 247	
Db	588	GVR 590	

RESULT 8  
US-09-418-540-6  
Invention 6, Application US/09418540  
Patent No. 6296848  
GENERAL INFORMATION:  
APPLICANT: POT, David A.  
APPLICANT: Williams, Lewis T.  
APPLICANT: Jefferson, Anne Bennett  
APPLICANT: Majerus, Philip W.  
TITLE OF INVENTION: No. 6296848el Grb2 Associating Protein and Nucleic  
Acids Encoding Therefor  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew  
STREET: One Market Plaza, Stewart tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/418,540

FILING DATE: 14-OCT-1999  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/560,005  
 FILING DATE: 17-NOV-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Dow, Karen B.  
 REGISTRATION NUMBER: 29,684  
 REFERENCE/DOCKET NUMBER: 2307K-0624000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-326-2400  
 TELEFAX: 415-326-2422  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 942 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 FEATURE:  
 NAME/KEY: Region  
 LOCATION: 1..942  
 OTHER INFORMATION: /note="majptase"  
 US-09-418-540-6

Query Match	21.4%	Score	432.5	DB	4	Length	942
Best Local Similarity	37.0%	Pred. No.	1.le-37				
Matches	90	Conservative	45	Mismatches	87	Indels	21
							Gaps
Qy	9	FIKSHVRMGOIILLVPKAYQHLPYIQLSTKTPGLFCYMGKNGKGVNICLKLYGYVVS	68				
Db	365	YAKVKLLRIGIMLLLVKQEHAAAYISEVEAETVGTGIMGRMCKGGVAIRFQFHTSIC	424				
Qy	69	IINCHLPPHISNNYQRLHF-DRILEMNCQCE-GRDIP--NILDHDLLIIWFGDMNFRIEDF	124				
Db	425	VVNSHLAAHIEEYERRNQDYKDCSRMFCQCPDPSLPPTISNHDVILMMGDLNYRIEEL	484				
Qy	125	GLHFVRESIKRNCYGGJLWEKQDLSIAKKHDPLLRETFQGRLLPEPTPYKFDNRNSNDYDTE	184				
Db	485	DVEKVKLLIEEKQFOMLYAYVDQLKIQVAAKTVPFEGTEGELTFQPTYKYDVTGSDDMWTS	544				
Qy	185	KKRKPANTDRILRLFKQPCAGDPPTPPASHFSLSLRGYSSHMTYGISDHKKPVSCTFDL	244				
Db	545	KCRAPACDRILMLKKGK-----NITQLSYQSHMALKTSDHKPVPSSVFDI	597				
Qy	245	ELK	247				
Db	588	GVR	590				

```

RESULT 9
US-08-560-005-7
; Sequence 7, Application US/08560005
; Patent No. 6001354
; GENERAL INFORMATION:
; APPLICANT: Pot, David A.
; APPLICANT: Williams, Lewis T.
; APPLICANT: Jefferson, Anne Bennett
; APPLICANT: Majerus, Philip W.
; TITLE OF INVENTION: No. 6001354el Grb2 Associating Protein and Nucleic
; TITLE OF INVENTION: Acids Encoding Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```







GENERAL INFORMATION:  
APPLICANT: Pot, David A.  
APPLICANT: Williams, Lewis T.  
APPLICANT: Jefferson, Anne Bennett  
APPLICANT: Majerus, Philip W.  
TITLE OF INVENTION: No. 6001354el Grb2 Associating Protein and Nucleic  
TITLE OF INVENTION: Acids Encoding Therefor  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/560,005  
FILING DATE: 14-OCT-1999  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 2307K-0624000  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..946  
OTHER INFORMATION: /note= "ysc5ptase"  
US-08-560-005-4

Query Match 14.4%; Score 290.5; DB 3; Length 946;  
Best Local Similarity 31.5%; Pred. No. 2.6e-22;  
Matches 90; Conservative 38; Mismatches 115; Indels 43; Gaps 8;  
Db 9 FIKVSHVRMOGILLVFAKYQHLPIYIQLSTKSTPTGLFGYWGNGGVNICLKLYGYVS 68  
607 YIRLWSTQLGGILLLLFMNETEYSKVKHIEGDVKKTFGCGMASNKGAVAVSFYSATRFC 666  
QY 69 IINCHLPPHISNNYQRLHEHFDRILE-MQNCGRDIPNLDHDLIIWFGDMNFR-----ED 123  
Db 667 VLVSHLAAGLENVEQRHNDYKTTAKSIRFSKGL---RIKDHDAIFWEGDENYRILASND 723  
QY 124 FGLHFVRESTKRCYCGGLWPKDOLS---IAKKHDPDLLREFQGRLLFPPTYKFDNRNDY 180  
Db 724 -----VRRKIVSKEYASLFEKDLQNOQMIAGESFPY---PHEMAIDFPPTYKFDNPKY 775  
QY 181 DTSEKKRKPATWDRILRLKRPQCAPDTPIPPASHFSLSLRGYSYSSMTYGISDHRPVSG 240  
Db 776 DTSEKMRPLPAWDRILSRGE-----VLEQLEYKCCEDILFSDHRPVYA 818  
QY 241 TF-----DLEKPLVSAPLIVLMPEDLWTVENDMMVSYSSSTDF 279  
Db 819 IFRARVTVDEQKKTTLGTQIYEKIMERLEGLDDDEKIAVLSDDAF 864

RESULT 14  
US-09-418-540-4  
; Sequence 4, Application US/09418540  
; Patent No. 6296848

GENERAL INFORMATION:  
APPLICANT: Pot, David A.  
APPLICANT: Williams, Lewis T.  
APPLICANT: Jefferson, Anne Bennett  
APPLICANT: Majerus, Philip W.  
TITLE OF INVENTION: No. 6296848el Grb2 Associating Protein and Nucleic  
TITLE OF INVENTION: Acids Encoding Therefor  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/418,540  
FILING DATE: 14-OCT-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/560,005  
FILING DATE: 17-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 2307K-0624000  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..946  
OTHER INFORMATION: /note= "ysc5ptase"  
US-09-418-540-4

Query Match 14.4%; Score 290.5; DB 4; Length 946;  
Best Local Similarity 31.5%; Pred. No. 2.6e-22;  
Matches 90; Conservative 38; Mismatches 115; Indels 43; Gaps 8;  
QY 9 FIKVSHVRMOGILLVFAKYQHLPIYIQLSTKSTPTGLFGYWGNGGVNICLKLYGYVS 68  
Db 607 YIRLWSTQLGGILLLLFMNETEYSKVKHIEGDVKKTFGCGMASNKGAVAVSFYSATRFC 666  
QY 69 IINCHLPPHISNNYQRLHEHFDRILE-MQNCGRDIPNLDHDLIIWFGDMNFR-----ED 123  
Db 667 VLVSHLAAGLENVEQRHNDYKTTAKSIRFSKGL---RIKDHDAIFWEGDENYRILASND 723  
QY 124 FGLHFVRESTKRCYCGGLWPKDOLS---IAKKHDPDLLREFQGRLLFPPTYKFDNRNDY 180  
Db 724 -----VRRKIVSKEYASLFEKDLQNOQMIAGESFPY---PHEMAIDFPPTYKFDNPKY 775  
QY 181 DTSEKKRKPATWDRILRLKRPQCAPDTPIPPASHFSLSLRGYSYSSMTYGISDHRPVSG 240  
Db 776 DTSEKMRPLPAWDRILSRGE-----VLEQLEYKCCEDILFSDHRPVYA 818  
QY 241 TF-----DLEKPLVSAPLIVLMPEDLWTVENDMMVSYSSSTDF 279  
Db 819 IFRARVTVDEQKKTTLGTQIYEKIMERLEGLDDDEKIAVLSDDAF 864

RESULT 15







C:Accession: S29069; I68621  
R:Attree, O.; Olivios, I.M.; Okabe, I.; Bailey, L.C.; Nelson, D.L.; Lewis, R.A.; McInnes, Nature 358, 239-242, 1992  
A:Title: The Lowe's oculocerebrorenal syndrome gene encodes a protein highly homologous A:Reference number: S29069; MUID:92334430  
A:Accession: S29069  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-970 <ATT>  
R:Leahy, A.M.; Charnas, L.R.; Nussbaum, R.L. Hum. Mol. Genet. 2, 461-463, 1993  
A:Title: Nonsense mutations in the OCRL-1 gene in patients with the oculocerebrorenal syndrome A:Reference number: I54349; MUID:93278398  
A:Accession: I68621  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 883-912 <RES>  
A:Cross-references: GB:S62085; NID:G385336; PIDN:AAB26926.1; PID:G385337  
C:Genetics:  
Gene: OCRL-1

Query Match 18.2%; Score 367.5; DB 2; Length 970;  
Best Local Similarity 29.4%; Pred. No. 1.3e-23;  
Matches 100; Conservative 58; Mismatches 131; Indels 51; Gaps 8;  
QY 5 SPLSFIKSHVRMOGILLVFAKYQHLPYIOILSTKSTPTGLFGYWGNGKGVNICKLYG 64  
Db 384 SKAKYKKVQLVRLVGMMLIFARKDCQRYIRDIATETVGTGKNGKNGKGVAVRFVHN 443  
QY 65 YVYSIINCHLPPHISNNYQRLHFHLEMON--CEGRDIP--NILDHDLIIWFGDMNFR 120  
Db 444 TTFCIVNSHLAAHVDFERNQDYKDICARMSFVNPQTLPLQINIMKHEVVIWGLDLYR 503  
QY 121 IEDFGLHFVRESIKNRCYGGIWEKDQLSIKKHDPDLRERQEGRLPPPTTKFDRNSNDY 180  
Db 504 LCPMDANEVKSILNKDLQRLKFDQLNQTQKAKAFVDFNEGEIKFIPYKYDSKTDWR 563  
QY 181 DTSEKKRKPAWTDRIILWLKRPQPCAGPDTPIPPASHFSLSLRGYSYSHMTYGISDHKPVSG 240  
Db 564 DSSGKCRVPAWCDRLWR-----GTNVNQLNYSHMLKTSDHKPVSA 606  
QY 241 TFDLELKPL-----VSAPLIVLMPEDLWTVENDMMVYS--STDFPSPWDWIGLYKV 292  
Db 607 LFHIGVKVVDERRYKRVFDSVRIMDR---MENDELPLSLSRREF-----649  
QY 293 GLRDVNDYVSANVGDSKVSCLNLAQVYIDISNIPTED 332  
650 ----VFENKFRLOLQKGFQISNN-QGVPCHFSPFIRPLND 684

RESULT 3  
A41075  
inositol-1,4,5-trisphosphate 5-phosphatase (EC 3.1.3.56) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 12-Jun-1992 #sequence\_revision 12-Jun-1992 #text\_change 29-Aug-1997  
C:Accession: A41075  
R:Ross, T.S.; Jefferson, A.B.; Mitchell, C.A.; Majerus, P.W. J. Biol. Chem. 266, 20283-20289, 1991  
A:Title: Cloning and expression of human 75-kDa inositol polyphosphate-5-phosphatase. A:Reference number: A41075; MUID:92041857  
A:Accession: A41075  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-672 <ROS>  
A:Cross-references: GB:M74161  
C:Genetics:  
A:Gene: GDB:INPP5B  
A:Cross-references: GDB:129756; OMIM:147264  
A:Map position: lp34-lp34  
C:Keywords: phosphoric monoester hydrolase

Query Match 16.5%; Score 334; DB 2; Length 672;  
Best Local Similarity 31.7%; Pred. No. 6.3e-21;  
Matches 78; Conservative 45; Mismatches 97; Indels 26; Gaps 6;  
QY 9 FIKVSHVRMOGILLVFAKYQHLPYIOILSTKSTPTGLFGYWGNGKGVNICKLYGYVS 68  
Db 133 YAKVKLIRLVGIMLLYVKEAAHYISEVAETVGTGIMGRMGKGVAFIRFQHNFSIC 192  
QY 69 IINCHLPPHISNNYQRLHF--DRILEMONCE--GRDIP--NILDHDLIIWFGDMNFRIDEF 124  
Db 193 VVNSHLAAHTEEYERNQDYKDCISRMQFCQDPDPSLPLTISNHDVILWLGDNLNRIEEL 252  
QY 125 GLHFVRESIKNRCYGGIWEKDQLSIKKHDPDLRERQEGRLPPPTTKFDRNSNDYDTSE 184  
Db 253 DVEKVKKLEEKDFQMLYAYDQLKIQAAKTVFEGFEGELTFQPTYKYDTRA--LTGTI 310  
QY 185 KKRKPWTDRI---LWLKRPQPCAGPDTPIPPASHFSLSLRGYSYSHMTYGISDHKPVSGT 241  
Db 311 PVRSAVLLPGVIGLWKGK-----NITQLSYQSHMALKTSDHKPVSSV 353  
QY 242 FDELEK 247  
Db 354 FDIGVR 359

RESULT 4  
JW0105  
synaptotjanin 2 alpha protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Jul-1998 #sequence\_revision 10-Jul-1998 #text\_change 08-Oct-1999  
C:Accession: JW0105  
R:Seet, L.F.; Cho, S.; Hessel, A.; Dumont, D.J. Biochem. Biophys. Res. Commun. 247, 116-122, 1998  
A:Title: Molecular cloning of multiple isoforms of synaptotjanin 2 and assignment of A:Reference number: JW0105; MUID:98300294  
A:Contents: Liver  
A:Accession: JW0105  
A:Molecule type: mRNA  
A:Residues: 1-1216 <SEE>  
A:Cross-references: GB:AF041862; NID:G3241994; PIDN:AAC40146.1; PID:G3241995  
Query Match 16.5%; Score 334; DB 2; Length 1216;  
Best Local Similarity 31.2%; Pred. No. 1.5e-20;  
Matches 92; Conservative 49; Mismatches 124; Indels 30; Gaps 9;  
QY 9 FIKVSHVRMOGILLVFAKYQHLPYIOILSTKSTPTGLFGYWGNGKGVNICKLYGYVS 68  
Db 536 YILLTSAQLVGVCLYIFVRPYHVPFIRDAIDTVKTMGMKAGNKGAVGIRFQLHSTSF 595  
QY 69 IINCHLPPHISNNYQRLHFDRIL-EMQNCGRDIPNILDHDLIIWFGDMNFRIDEFGLH 127  
Db 596 FVCSHLTAGOSQVKERNEDYREITHKLSPSGR--NIFSHDYVFWCGDFNYRI-DLTYE 651  
QY 128 FVRESIKNRCYGGIWEKDQLSIKKHDPDLRERQEGRLPPPTTKFDRNSNDYDTSEKKR 187  
Db 652 EVFYFKRODWKLMFEFDQLQKSSGKIFKDFHEGAVNFGPTYKYDVGSAAYDTSKCR 711  
QY 188 KPAWTDRIILWLKRPQ---CAGP---DTPI---PPASH--FSLSLRGYSYSHMTYGISDHK 236  
Db 712 TPWTDRIILWLKRPQ---CAGP---DTPI---PPASH--FSLSLRGYSYSHMTYGISDHK 236  
QY 237 PVSGTFDELEKPL-----VSAPLIVLMPEDLWTVENDMMVYSYSTSDFP 280  
Db 772 PVLAIVEVEVQVDVGARERVFQEVSS---VOGPLDATVVVNQLQSPLEEKNEFP 823

RESULT 5  
S68448  
synaptotjanin, 170K - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 05-Nov-1999  
C:Accession: S68448; S78547; S78527

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	432.5	21.4	942	1	ISP2_HUMAN	IP32019	homo sapien
2	367.5	18.2	901	1	OCRL_HUMAN	Q01968	homo sapien
3	328	16.2	1574	1	SVJ1_RAT	O62910	rattus norv
4	328	16.2	1575	1	SVJ1_HUMAN	O43426	homo sapien
5	327.5	16.2	1443	1	SVJ2_HUMAN	O15056	homo sapien
6	326	16.1	1248	1	SVJ2_RAT	O55207	rattus norv
7	325	16.1	1324	1	SVJ1_BOVIN	O18964	bos taurus
8	289.5	14.3	946	1	YTA2_YEAST	P50942	saccharomyc
9	265	13.1	1183	1	YNK6_YEAST	P40559	saccharomyc
10	227.5	11.3	398	1	YJL7_CAEEL	P34370	caenorhabdi
11	142.5	7.1	396	1	YSP1_CAEEL	Q17848	caenorhabdi
12	114	5.6	743	1	YCR6_YEAST	P23533	saccharomyc
13	104.5	5.2	412	1	ISP1_HUMAN	Q14642	homo sapien
14	102.5	5.1	412	1	ISP1_CANFA	Q29467	canis famill
15	99	4.9	423	1	NSMA_HUMAN	O60906	homo sapien
16	95.5	4.7	954	1	YB79_YEAST	P38138	saccharomyc
17	93	4.6	389	1	YP46_STBVP	P27501	rice tungro
18	93	4.6	1416	1	BLM_MOUSE	O88700	mus musculus
19	91	4.5	1070	1	AGLU_CANTS	P29064	candida tsu
20	87.5	4.3	475	1	MM03_RAT	P03957	rattus norv
21	87	4.3	338	1	PRSC_DROME	P26270	drosophila
22	87	4.3	626	1	HCYB_EURCA	O9nf9h	eurytelma c
23	86.5	4.3	251	1	Y761_METJA	Q58171	methanococc
24	86	4.3	447	1	GN22_HUMAN	Q10469	h alpha-1,6
25	86	4.3	2034	1	FER1_CAEEL	Q13428	caenorhabdi
26	85	4.2	764	1	PAG_BACAN	P13783	bacillus an
27	85	4.2	4644	1	DYHC_RAT	P38650	rattus norv
28	84.5	4.2	566	1	CPYC_EMENI	P87208	emericeella
29	84.5	4.2	993	1	EPB3_MOUSE	P54754	mus musculu
30	84	4.2	402	1	NSNA_DROME	Q9vzs6	drosophila
31	84	4.2	442	1	GN22_RAT	O09326	r alpha-1,6
32	83.5	4.1	515	1	STF7_YEAST	P06784	saccharomyc
33	83.5	4.1	552	1	REFP_FOWPV	O72909	fowlpox vir





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 09:21:54 ; Search time 28.13 Seconds  
(without alignments)  
2287.740 Million cell updates/sec

Title: US-09-892-287-1  
Perfect score: 2020  
Sequence: 1 MDVLSPLSFVKVSHVRMGI.....FQIPPGSLREDPLGEAQPOI 372

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Maximum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp Vertebrate.\*
- 14: sp Unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2020	100.0	448	Q9BT40	Q9bt40 homo sapien
2	2010	99.5	372	Q9P2R5	Q9p2r5 homo sapien
3	2010	99.5	448	Q9NPJ5	Q9npj5 homo sapien
4	1791	88.7	329	Q15733	Q15733 homo sapien
5	1525	75.5	468	O09040	O09040 mus musculus
6	872.5	43.2	1001	Q9JMC1	Q9jmc1 rattus norv
7	831.5	41.2	1056	Q9UDT9	Q9udt9 homo sapien
8	507.5	25.1	508	Q9VXE7	Q9vxe7 drosophila
9	500.5	24.8	397	Q15735	Q15735 homo sapien
10	460.5	22.8	357	Q9V7X0	Q9v7x0 drosophila
11	423.5	21.0	749	Q9DBW2	Q9dbw2 mus musculus
12	423.5	21.0	776	O54996	O54996 mus musculus
13	423.5	21.0	825	O151Z8	O151z8 mus musculus
14	340	16.8	1113	Q9XUD3	Q9xud3 caenorhabdi
15	340	16.8	1119	O9GT42	O9gt42 caenorhabdi
16	339	16.8	1291	Q9DGN7	Q9dgn7 lampetra fl

17	338	16.7	1101	10	Q9XFT1	Q9xft1 arabidopsis
18	334	16.5	824	11	Q9D2G5	Q9d2g5 mus musculu
19	334	16.5	1479	11	O88399	O88399 mus musculu
20	328	16.2	1193	4	O94984	O94984 homo sapien
21	327.5	16.2	1496	4	Q9H226	Q9h226 homo sapien
22	326	16.1	1451	11	Q91ZD9	Q91zd9 rattus norv
23	326	16.1	1496	11	Q91ZD8	Q91zd8 rattus norv
24	316.5	15.7	1218	5	Q9W296	Q9w296 drosophila
25	309.5	15.3	1136	10	Q9SYK4	Q9syk4 arabidopsis
26	306	15.1	1305	10	O80560	O80560 arabidopsis
27	305.5	15.1	609	10	Q9LR47	Q9lr47 arabidopsis
28	304.5	15.1	1171	3	Q96U77	Q96u77 neurospora
29	303.5	15.0	850	5	O46094	O46094 drosophila
30	300.5	14.9	753	5	O17590	O17590 caenorhabdi
31	298.5	14.8	331	10	Q9SX81	Q9sx81 arabidopsis
32	297.5	14.7	401	10	Q9ZUS3	Q9zus3 arabidopsis
33	291.5	14.4	1107	3	Q12271	Q12271 saccharomyc
34	284	14.1	1144	10	Q9SKB7	Q9skb7 arabidopsis
35	284	14.1	1149	4	Q13577	Q13577 homo sapien
36	284	14.1	1258	4	O15357	O15357 homo sapien
37	282	14.0	1183	11	Q9R1V2	Q9riv2 rattus norv
38	282	14.0	1257	11	Q9WVR3	Q9wvr3 rattus norv
39	281.5	13.9	321	5	Q95041	Q95041 rhodnius pr
40	281	13.9	1257	11	Q9JLL7	Q9jll7 mus musculu
41	279.5	13.8	1076	3	O43001	O43001 schizosacch
42	279	13.8	501	10	Q9SKZ8	Q9skz8 arabidopsis
43	278.5	13.8	586	10	Q9FUR3	Q9fur3 arabidopsis
44	278.5	13.8	590	10	Q9FX20	Q9fx20 arabidopsis
45	278.5	13.8	1191	3	O14306	O14306 schizosacch

ALIGNMENTS

RESULT 1

Q9BT40	ID	Q9BT40	PRELIMINARY;	PRT;	448 AA.
AC	Q9BT40;				
DT	01-JUN-2001	(Tfemblrel. 17, Created)			
DT	01-JUN-2001	(Tfemblrel. 17, Last sequence update)			
DT	01-DEC-2001	(Tfemblrel. 19, Last annotation update)			
DE	SKIP FOR SKELETAL MUSCLE AND KIDNEY ENRICHED INOSITOL				
DE	PHOSPHATASE.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=EYE, RETINOBLASTOMA;				
RA	Strausberg R.;				
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC004362; AA04362.1;				
DR	InterPro; IPR000300; IPPC.				
DR	Pfam; PF00783; IPPC; 1.				
DR	SMART; SM00128; IPPC; 1.				
SQ	SEQUENCE 448 AA; 51090 MW; 46FAA48C6E2EAD4 CRC64;				
Query Match 100.0%; Score 2020; DB 4; Length 448;					
Best Local Similarity 100.0%; Pred. No. 1.7e-178;					
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MDVLSPLSFVKVSHVRMGI			60
Db	77	MDVLSPLSFVKVSHVRMGI			136
QY	61	KLGYVYVSIINCHLPPHISNNYORLEHFDHFRILEMNCNCEGRDIPNILDHDLIIWFGDMNFR			120
Db	137	KLGYVYVSIINCHLPPHISNNYORLEHFDHFRILEMNCNCEGRDIPNILDHDLIIWFGDMNFR			196
QY	121	IEDFGLHFVRESIKNRCYGGIWEKDQLSIAKKHDPILLREFEGGELLFPPTPKKDRNSNDY			180

```
Db 197 IEDFGLHFVRESIKNRCYGLWEKDQLSIKAKHDPDLLREFQEGRLLPPTTKFDRNSNDY 256
QY 181 DTSEKKRKPAWTDRIILWRLKQPCAGDTPIPPASHFSLSLRGYSSTMHTYGISDHPKPVSG 240
Db 257 DTSEKKRKPAWTDRIILWRLKQPCAGDTPIPPASHFSLSLRGYSSTMHTYGISDHPKPVSG 316
QY 241 TFDELEKPLVSAPLIVLMPEDLMTVNDMMVSYSSDTPSPDWMTGLYKVGRLDVNDY 300
Db 317 TFDELEKPLVSAPLIVLMPEDLMTVNDMMVSYSSDTPSPDWMTGLYKVGRLDVNDY 376
QY 301 VSYAWGDSKVSCSDNLNQYIIDISNPTTEDEFLCYNSLSRVVGVISRRPFIIPGSL 360
Db 377 VSYAWGDSKVSCSDNLNQYIIDISNPTTEDEFLCYNSLSRVVGVISRRPFIIPGSL 436
QY 361 REDPLGEAQPOI 372
Db 437 REDPLGEAQPOI 448

RESULT 2
ID Q9NPJ5 PRELIMINARY; PRT; 372 AA.
AC Q9NPJ5;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DE 43-KDA FORM SKELETAL MUSCLE AND KIDNEY ENRICHED INOSITOL
DE PHOSPHATASE.
GN SKIP 43-KDA FORM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ijuin T., Mochizuki Y., Fukami K., Funaki M., Asano T., Takenawa T.;
RT Identification and Characterization of a Novel Inositol Polyphosphate
RT 5-Phosphatase.
RL J. Biol. Chem. 0:0-0(2000).
DR EMBL; AB036831; BAA92341.1; -.
DR InterPro; IPR000300; IPPC.
DR Pfam; PF00783; IPPC; 1.
DR SMART; SM00128; IPPC; 1.
SQ SEQUENCE 372 AA; 42922 MW; A2FCAE390DCD4528 CRC64;

Query Match 99.5%; Score 2010; DB 4; Length 372;
Best Local Similarity 99.5%; Pred. No. 1.1e-177;
Matches 370; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDVLSPLSFVKVSHVRMQGILLVFAKYQHLPIYIQLSTKSTPTGLFGYGNKGGVNICL 60
Db 1 MDVLSPLSFVKVSHVRMQGILLVFAKYQHLPIYIQLSTKSTPTGLFGYGNKGGVNICL 60
QY 61 KLYGYVSIINCHLPHISNNYORLEHFDRILEMQCEGRDIPNILDHDLIIWFGDMNFR 120
Db 61 KLYGYVSIINCHLPHISNNYORLEHFDRILEMQCEGRDIPNILDHDLIIWFGDMNFR 120
QY 121 IEDFGLHFVRESIKNRCYGLWEKDQLSIKAKHDPDLLREFQEGRLLPPTTKFDRNSNDY 180
Db 121 IEDFGLHFVRESIKNRCYGLWEKDQLSIKAKHDPDLLREFQEGRLLPPTTKFDRNSNDY 180
QY 181 DTSEKKRKPAWTDRIILWRLKQPCAGDTPIPPASHFSLSLRGYSSTMHTYGISDHPKPVSG 240
Db 181 DTSEKKRKPAWTDRIILWRLKQPCAGDTPIPPASHFSLSLRGYSSTMHTYGISDHPKPVSG 240
QY 241 TFDELEKPLVSAPLIVLMPEDLMTVNDMMVSYSSDTPSPDWMTGLYKVGRLDVNDY 300
Db 241 TFDELEKPLVSAPLIVLMPEDLMTVNDMMVSYSSDTPSPDWMTGLYKVGRLDVNDY 300
QY 301 VSYAWGDSKVSCSDNLNQYIIDISNPTTEDEFLCYNSLSRVVGVISRRPFIIPGSL 360
Db 301 VSYAWGDSKVSCSDNLNQYIIDISNPTTEDEFLCYNSLSRVVGVISRRPFIIPGSL 360
QY 361 REDPLGEAQPOI 372
Db 437 REDPLGEAQPOI 448

RESULT 3
ID Q9NPJ5 PRELIMINARY; PRT; 448 AA.
AC Q9NPJ5;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE SKELETAL MUSCLE AND KIDNEY ENRICHED INOSITOL PHOSPHATASE.
GN SKIP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ijuin T., Mochizuki Y., Fukami K., Funaki M., Asano T., Takenawa T.;
RT Identification and Characterization of a Novel Inositol Polyphosphate
RT 5-Phosphatase.
RL J. Biol. Chem. 0:0-0(2000).
DR EMBL; AB036831; BAA92341.1; -.
DR InterPro; IPR000300; IPPC.
DR Pfam; PF00783; IPPC; 1.
DR SMART; SM00128; IPPC; 1.
SQ SEQUENCE 448 AA; 51228 MW; 0651BE3C6E2EEB0F CRC64;

Query Match 99.5%; Score 2010; DB 4; Length 448;
Best Local Similarity 99.5%; Pred. No. 1.4e-177;
Matches 370; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDVLSPLSFVKVSHVRMQGILLVFAKYQHLPIYIQLSTKSTPTGLFGYGNKGGVNICL 60
Db 77 MDVLSPLSFVKVSHVRMQGILLVFAKYQHLPIYIQLSTKSTPTGLFGYGNKGGVNICL 136
QY 61 KLYGYVSIINCHLPHISNNYORLEHFDRILEMQCEGRDIPNILDHDLIIWFGDMNFR 120
Db 137 KLYGYVSIINCHLPHISNNYORLEHFDRILEMQCEGRDIPNILDHDLIIWFGDMNFR 196
QY 121 IEDFGLHFVRESIKNRCYGLWEKDQLSIKAKHDPDLLREFQEGRLLPPTTKFDRNSNDY 180
Db 197 IEDFGLHFVRESIKNRCYGLWEKDQLSIKAKHDPDLLREFQEGRLLPPTTKFDRNSNDY 256
QY 181 DTSEKKRKPAWTDRIILWRLKQPCAGDTPIPPASHFSLSLRGYSSTMHTYGISDHPKPVSG 240
Db 257 DTSEKKRKPAWTDRIILWRLKQPCAGDTPIPPASHFSLSLRGYSSTMHTYGISDHPKPVSG 316
QY 241 TFDELEKPLVSAPLIVLMPEDLMTVNDMMVSYSSDTPSPDWMTGLYKVGRLDVNDY 300
Db 317 TFDELEKPLVSAPLIVLMPEDLMTVNDMMVSYSSDTPSPDWMTGLYKVGRLDVNDY 376
QY 301 VSYAWGDSKVSCSDNLNQYIIDISNPTTEDEFLCYNSLSRVVGVISRRPFIIPGSL 360
Db 377 VSYAWGDSKVSCSDNLNQYIIDISNPTTEDEFLCYNSLSRVVGVISRRPFIIPGSL 436
QY 361 REDPLGEAQPOI 372
Db 437 REDPLGEAQPOI 448

RESULT 4
ID Q15733 PRELIMINARY; PRT; 329 AA.
AC Q15733;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
```

```
Db 197 IEDFGLHFVRESIKNRCYGLWEKDQLSIKAKHDPDLLREFQEGRLLPPTTKFDRNSNDY 256
QY 181 DTSEKKRKPAWTDRIILWRLKQPCAGDTPIPPASHFSLSLRGYSSTMHTYGISDHPKPVSG 240
Db 257 DTSEKKRKPAWTDRIILWRLKQPCAGDTPIPPASHFSLSLRGYSSTMHTYGISDHPKPVSG 316
QY 241 TFDELEKPLVSAPLIVLMPEDLMTVNDMMVSYSSDTPSPDWMTGLYKVGRLDVNDY 300
Db 317 TFDELEKPLVSAPLIVLMPEDLMTVNDMMVSYSSDTPSPDWMTGLYKVGRLDVNDY 376
QY 301 VSYAWGDSKVSCSDNLNQYIIDISNPTTEDEFLCYNSLSRVVGVISRRPFIIPGSL 360
Db 377 VSYAWGDSKVSCSDNLNQYIIDISNPTTEDEFLCYNSLSRVVGVISRRPFIIPGSL 436
QY 361 REDPLGEAQPOI 372
Db 437 REDPLGEAQPOI 448

RESULT 2
ID Q9P2R5 PRELIMINARY; PRT; 372 AA.
AC Q9P2R5;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DE 43-KDA FORM SKELETAL MUSCLE AND KIDNEY ENRICHED INOSITOL
DE PHOSPHATASE.
GN SKIP 43-KDA FORM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ijuin T., Mochizuki Y., Fukami K., Funaki M., Asano T., Takenawa T.;
RT Identification and Characterization of a Novel Inositol Polyphosphate
RT 5-Phosphatase.
RL J. Biol. Chem. 0:0-0(2000).
DR EMBL; AB036830; BAA92341.1; -.
DR InterPro; IPR000300; IPPC.
DR Pfam; PF00783; IPPC; 1.
DR SMART; SM00128; IPPC; 1.
SQ SEQUENCE 372 AA; 42922 MW; A2FCAE390DCD4528 CRC64;

Query Match 99.5%; Score 2010; DB 4; Length 372;
Best Local Similarity 99.5%; Pred. No. 1.1e-177;
Matches 370; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDVLSPLSFVKVSHVRMQGILLVFAKYQHLPIYIQLSTKSTPTGLFGYGNKGGVNICL 60
Db 1 MDVLSPLSFVKVSHVRMQGILLVFAKYQHLPIYIQLSTKSTPTGLFGYGNKGGVNICL 60
QY 61 KLYGYVSIINCHLPHISNNYORLEHFDRILEMQCEGRDIPNILDHDLIIWFGDMNFR 120
Db 61 KLYGYVSIINCHLPHISNNYORLEHFDRILEMQCEGRDIPNILDHDLIIWFGDMNFR 120
QY 121 IEDFGLHFVRESIKNRCYGLWEKDQLSIKAKHDPDLLREFQEGRLLPPTTKFDRNSNDY 180
Db 121 IEDFGLHFVRESIKNRCYGLWEKDQLSIKAKHDPDLLREFQEGRLLPPTTKFDRNSNDY 180
QY 181 DTSEKKRKPAWTDRIILWRLKQPCAGDTPIPPASHFSLSLRGYSSTMHTYGISDHPKPVSG 240
Db 181 DTSEKKRKPAWTDRIILWRLKQPCAGDTPIPPASHFSLSLRGYSSTMHTYGISDHPKPVSG 240
QY 241 TFDELEKPLVSAPLIVLMPEDLMTVNDMMVSYSSDTPSPDWMTGLYKVGRLDVNDY 300
Db 241 TFDELEKPLVSAPLIVLMPEDLMTVNDMMVSYSSDTPSPDWMTGLYKVGRLDVNDY 300
QY 301 VSYAWGDSKVSCSDNLNQYIIDISNPTTEDEFLCYNSLSRVVGVISRRPFIIPGSL 360
Db 301 VSYAWGDSKVSCSDNLNQYIIDISNPTTEDEFLCYNSLSRVVGVISRRPFIIPGSL 360
```

DR Pfam: PF00783; IPpe; 1.  
DR SMART: SM00128; IPPC; 1.  
SQ SEQUENCE 468 AA; 54158 MW; F2E1CA370B97A8A1 CRC64;

Query Match 75.5%; Score 1525; DB 11; Length 468;  
Best Local Similarity 75.7%; Pred. No. 1.2e-132;  
Matches 283; Conservative 39; Mismatches 50; Indels 2; Gaps

QY 1 MDVLSPLSFITKVSRRVMOGILLLVFAKYQHLPYIIILSTKSTPTGLFGYWGNGKGVNCL 60  
Db 95 MDMLSPNFVKISQVRMOGILLLVFAKYQHLPYIIILSTKSTPTGLFGYWGNGKGVNCL 154

QY 61 KLYGYVSIINCHLPHPHISNNYORLEHFDHLEQNGEGRDIPNILDHDLIIWFGDMNFR 120  
Db 155 KLYGYVSIINCHLPHPHMYNNDQRLHFDHLESLTTEFGYDVPNILDHDLIIWFGDMNFR 214

QY 121 IEDFGHLHFVRESITKNCRYCGGLWEKDQLSIAKKHDPDLREFQEGRLLPPTYKFFDRNSNDY 180  
Db 215 IEDFGLLFVQESITRKYKELWEKDQLFIAKNDQLLREFQEGRLLPPTYKFFDRHSNNY 274

QY 181 DTSEKKRKPATWDRILWRLKRQPC-AGDPDTPPPASHFSLSRLQYSSHMTYGISDHPKVS 239  
Db 275 DTSEKKRKPATWDRILWRLKRQPSQASPLASSVPTSYFLTLTKNVSHMAYISDHPKPT 334

QY 240 GTFDELEKPLVASPLVILMPEDLWTVENDMMVSYSSDPPSPQDWIGLYKVGLRVND 299  
Db 335 GTFDELELPLMSVPLITMPPEHLWTMDMLISTPTEFLSSWDWIGLYKVGMRHND 394

QY 300 YVYAWYGDSKSCDNLNQYIIDISNIPTEDEFLCYYSNSRVSVVGISRPFQIPPGS 359  
Db 395 YVAYVWVGDNQVSGNPNQVYINSAIPDTEQDFLLCYYSNNLSHVVGISQPFKIPRS 454

QY 360 -LREDPLGEAQPOI 372  
Db 455 FLREDTLYEPEPOI 468

RESULT 6  
Q9JMC1  
ID Q9JMC1 PRELIMINARY; PRT; 1001 AA.

AC Q9JMC1;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE PROLINE-RICH INOSITOL POLYPHOSPHATE 5-PHOSPHATASE.  
GN PIP.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
ON NCBI\_TaxID=10116;  
RX [1]  
RN SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RP MEDLINE=20062911; PubMed=10593988;  
RX Medchizuki Y.; Takenawa T.;  
RA "Novel inositol polyphosphate 5-phosphatase localizes at membrane  
RT ruffles";  
RL J. Biol. Chem. 274:36790-36795(1999).  
RL EMBL; AB032551; BAA90553.1; -;  
DR InterPro; IPR000300; IPPC.  
DR InterPro; IPR002965; P\_rich\_extensn.  
DR Pfam; PF00783; IPPC; 1.  
DR PRINTS; PR01217; PRICHEXTENS.  
DR SMART; SM00128; IPPC; 1.  
SQ SEQUENCE 1001 AA; 107207 MW; 7BE7741FEF8F3FAB CRC64;

Query Match 43.2%; Score 872.5; DB 11; Length 1001;  
Best Local Similarity 47.5%; Pred. No. 7.8e-72;  
Matches 169; Conservative 56; Mismatches 126; Indels 5; Gaps

OV 1 MDVLSPLSFITKVSRRVMOGILLLVFAKYQHLPYIIILSTKSTPTGLFGYWGNGKGVNCL 60

Db	483	MDALGPENFVLSVTRMQGVILLFKAYIHLPELRDQVDC	TFTGLGGVWGNKGVS	542
Qy	61	KLXYGVYSIINCLPHPRISNNYQRLSEHFDRILEM	QCEGRDIPNILDHLLIITWFGDMNR	120
Db	543	AATGHMLCFLNCHLPAHMDKAEQKDNFTILSQ	QFGGAHGLLDHDLVFWFGDLNFR	602
Qy	121	IEDFGHLHFVRESITKNRCYGGLEWKEDOLSI	AKKHDPLLREFQGRGLFPPTYKFRDMSNDY	180
Db	603	IESYDILHFVKFAIDSQHLQLEWKEDQLNMAKNT	WPTLKGFOGPLNFAFTFKFQVDGVTNKY	662
Qy	181	DTSEKKRKPANTORILWLRLKQPCAGDPDPI	PPASH-FSLSLRGVYSSHMTYIGISDHKKPVS	239
Db	663	DTSAKKRKPANTORILWLKVK-APSGGP-SP	SGRESHRLQVTOHYSKSHMEYTVSDHKPVA	720
Qy	240	GTDLLELKLPLVSAPLVLMPEDLWTVENDMV	SYSTSDFPSGSPMDWIGLYKVGGLRDND	299
Db	721	ARELLOFAFRDDVPLVRLEVADEWARPEQAV	RYRYETVFASSDWIDWIGLYRVGFHCKD	780
Db	300	YVSVMWGDSKVSCDNLNQVYIDISNIPTTE	DFELLCYYYSNLSRVVGISRPFOI	355
Db	781	XVAYVWAKHEV--DGNIVQVTSSESLPKGHG	DFILGYVSHHHSILLIGVTEPFOI	834

```

RESULT      7
Q9UDT9
ID          Q9UDT9      PRELIMINARY;      PRT;      1056 AA.
AC          Q9UDT9
DT          01-MAY-2000 (TrEMBLrel. 13, Created)
DT          01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT          01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE          WOGSC:H_DJ412A9.2 PROTEIN (FRAGMENT).
DE          WOGSC:H_DJ412A9.2.
GN          Homo sapiens (Human).
OS          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX          NCBI_TaxId=9606;
RN          [1]
RP          SEQUENCE FROM N.A.
RX          MEDLINE=99063792; PubMed=9847074;
RA          Sulston J.E., Waterston R.;
RT          "Toward a complete human genome sequence.";
RL          Genome Res. 8:1097-1108(1998).
RN          [2]
RP          SEQUENCE FROM N.A.
RX          Murray J., Lennox S., Harmon G.;
RT          "The sequence of Homo sapiens PAC clone RP3-412A9.";
RL          Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN          [3]
RP          SEQUENCE FROM N.A.
RX          Waterston R.;
RL          Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RL          EMBL; AC005005; AAD15618.1; -.
DR          InterPro; IPR000300; IPPC.
DR          InterPro; IPR002965; P_rich_extensn.
DR          Pfam; PF00783; IPPC; 2.
DR          PRINTS; PR01217; PRICHTEXTENS.
DR          SMART; SM00128; IPPC; 1.
SQ          NON_TER      1
          SEQUENCE      1056 AA; 112564 MW; B82A8C134D727D0 CRC64;

```

QY	87	HFRDILEMNCCEGRDIPNILDHD----	LIIWFGDMNFRIEDFGLHFRVRSIKNRCCGGWL	14
Dd	620	NFOITILSLQOFGQGAQIUDHEYGVLGFWFEGDLNFRIESYDLHFVKFAIDSDQLHQWL	679	
QY	143	EKQOLSTAKKHDPDLLREFOBGRLLFPPTTKYKFDNRNSNDYDTSEKKRKPANTDRILWLKRQ	202	
Dd	680	EKQOLNAKWNTWILKFQGBEPLNFAPTFKFDVGTNKYDTSAKRRKPANTDRILWKVK-A	738	
QY	203	PCAGPOTPIPPASH-FSLSLRGYSHHWTYGISDHKPKVSGTDFLELKLPLVASAPLIVLMPE	261	
Dd	739	PGGP-SPSGRKSHRLQVTQHSYRSMEYTVSDHKPVAAQFLIQAFARDMDPLVRLEVAD	797	
QY	262	LWTVENDMMVSYSTSDFPSPPDWMTGLYKVGLRDVNDYVSVAWVGDSKVCSDNLNOVY	321	
Dd	798	EWVRPEOAUVRYRMETVFARSSWDWTLGVRFGRHCXDVAYVMAKHED--DGNTYQT	855	
QY	322	IDISNTPTEDEFLCYYSNLSRVVGVISRPFOI	355	
Dd	856	FSEELPKGHGDFFILGYSHNHILIGITEPFOI	889	
RESULT	8			
Q9VXE7	IC	Q9VXE7	PRELIMINARY;	PRT; 508 AA.
AD	CQVXE7			
DT	01-MAY-2000	(TREMBLrel_13, Created)		
DT	01-MAY-2000	(TREMBLrel_13, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel_17, Last annotation update)		
DN	CG9784	PROTEIN.		
GN	CG9784.			
OS	Drosophila melanogaster	(Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Prterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxId=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RC	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriel J.F., Abayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Balaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Berson K.Y., Beron P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J.J., Brokstein P., Brotter P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,			
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,			
RA	Pallazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Swirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,			
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: August 6, 2002, 09:17:23 ; Search time 31.79 Seconds  
(without alignments)  
1299.763 Million cell updates/sec

Title: US-09-892-287-1  
Perfect score: 2020  
Sequence: 1 MDVLSPLFKVSHVMQGI.....FQIPPGSLREDPLGEAOFQI 372

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Top number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
1:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:	*	
2:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:	*	
3:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:	*	
4:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:	*	
5:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:	*	
6:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:	*	
7:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:	*	
8:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:	*	
9:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:	*	
10:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:	*	
11:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:	*	
12:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:	*	
13:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:	*	
14:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:	*	
15:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:	*	
16:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:	*	
17:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:	*	
18:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:	*	
19:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:	*	
20:	/SID55/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:	*	
21:	/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:	*	
22:	/SID55/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:	*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	2020	100.0 372 20	AAW97094 Phosphatidylinositol
2	683.5	33.8 279 22	AAE73981 Human colon cancer
3	637.5	31.6 255 21	AAE27797 Human secreted pro
4	507.5	25.1 508 22	ABE64662 Drosophila melanog
5	494.5	24.5 381 21	AAE27845 Sequence homology
6	460.5	22.8 357 22	ABE62704 Drosophila melanog
7	432.5	21.4 749 22	ABE95181 Human protein sequ
8	428	21.2 382 21	AAE27846 Protein fragment e
9	316.5	15.7 1218 22	ABE62412 Drosophila melanog
10	316.5	15.7 1218 22	ABE67376 Drosophila melanog
11	303.5	15.0 850 22	ABE59981 Drosophila melanog

12	284	14.1	1258	21	AAW80120	Human Ship-2 prote
13	284	14.1	1258	22	AAE98987	Human type 2 SH2-d
14	252	12.5	968	18	AAW21721	SH2-A. Homo sapie
15	251	12.4	1187	18	AAW14002	Mouse SH2-containi
16	250	12.4	976	18	AAW26624	Signalling inosito
17	250	12.4	976	18	AAW18327	BR2 associating p
18	250	12.4	1187	18	AAW14003	Human SH2-containi
19	250	12.4	1189	18	AAW26623	Signalling inosito
20	244.5	12.1	747	22	ABE71929	Drosophila melanog
21	225.5	11.2	141	22	AAU23513	Novel human enzyme
22	205	10.1	188	22	AAE85356	Human phosphatase
23	117.5	5.8	708	22	ABE62748	Drosophila melanog
24	104.5	5.2	458	22	ABE26947	Novel human diagno
25	101	5.0	539	21	AAE10224	Chicken neurite ei
26	99	4.9	423	20	AAW93359	WO 99/07855 SeqID
27	99	4.9	423	22	AAW40244	Human polypeptide
28	98.5	4.9	542	22	ABE63962	Drosophila melanog
29	98	4.9	314	22	AAW40245	Human polypeptide
30	95	4.7	433	22	AAW42031	Human polypeptide
31	92	4.6	433	22	AAW42030	Human polypeptide
32	91	4.5	580	22	AAW42158	Human polypeptide
33	91	4.5	580	22	AAW42159	Human polypeptide
34	91	4.5	690	22	AAE95603	Human protein sequ
35	91	4.5	691	22	AAW40372	Human polypeptide
36	91	4.5	691	22	AAU12179	Human PRO4996 poly
37	91	4.5	1070	13	AAE21521	Alpha galactosidas
38	90.5	4.5	497	22	ABG03007	Novel human diagno
39	89	4.4	1048	22	ABG21867	Novel human diagno
40	88	4.4	551	22	ABG20239	Novel human diagno
41	87.5	4.3	106	21	AAE01453	Human secreted pro
42	87.5	4.3	463	10	AAE93630	Sequence of rat tr
43	87.5	4.3	691	14	AAE38735	Beta-galactosidase
44	87	4.3	735	15	AAE60179	Protective antigen
45	87	4.3	736	21	AAE56959	B. anthracis MAT-p

ALIGNMENTS

RESULT 1						
AAW97094						
ID	AAW97094	standard; Protein; 372 AA.				
XX						
AC	AAW97094;					
XX						
DT	28-APR-1999	(first entry)				
XX						
DE	Phosphatidylinositol 4, 5-bisphosphate 5-phosphatase.					
XX						
KW	Human; phosphatidylinositol 4,5-bisphosphate 5-phosphatase; PBPP;					
KW	Incyte clone 638789; antagonist; immune disorder; cancer;					
KW	neuronal disorder; human tubby homologue.					
XX						
OS	Homo sapiens.					
XX						
FH	Key	Location/Qualifiers				
FT	Modified-site	38				
FT	Modified-site	/note= "potential phosphorylation site"				
FT	Modified-site	132				
FT	Modified-site	/note= "potential phosphorylation site"				
FT	Modified-site	170				
FT	Modified-site	/note= "potential phosphorylation site"				
FT	Modified-site	183				
FT	Modified-site	/note= "potential phosphorylation site"				
FT	Modified-site	192				
FT	Modified-site	/note= "potential phosphorylation site"				
FT	Modified-site	275				
FT	Modified-site	/note= "potential phosphorylation site"				
FT	Modified-site	282				
FT	Modified-site	/note= "potential phosphorylation site"				
FT	Modified-site	295				
FT	Modified-site	/note= "potential phosphorylation site"				
FT	Modified-site	312				

FT Modified-site /note= "potential phosphorylation site"  
FT 329  
FT Modified-site /note= "potential phosphorylation site"  
FT 330  
FT Modified-site /note= "potential phosphorylation site"  
FT 339  
FT Modified-site /note= "potential phosphorylation site"  
FT 359  
XX W09900507-A1.  
XX  
XX  
XX PD 07-JAN-1999.  
XX  
XX PF 26-JUN-1998; 98WO-US13399.  
XX PR 27-JUN-1997; 97US-0884681.  
XX  
XX (INCY-) INCYTE PHARM INC.  
XX  
XX Corley NC, Hillman JL, Lal P, Shah P;  
BT WPI; 1999-095752/08.  
DR N-PSDB; AAX15254.  
XX  
XX Phosphatidylinositol 4,5-bisphosphate 5-phosphatase - used for  
PT treating immune disorders, cancers, and neuronal disorders  
PT  
XX  
XX PS Claim 1; Fig 1A-G; 81pp; English.  
XX  
XX The present sequence encodes a human phosphatidylinositol  
CC 4,5-bisphosphate 5-phosphatase (PBP) protein. The PBP polynucleotide  
CC was first identified in Incyte clone 638789 from the breast cancer  
CC cDNA library BRSTNOT03. Antagonists of the PBP protein can be used  
CC in the treatment or prevention of an immune disorder, a cancer, or  
CC a neuronal disorder. The PBP polynucleotide can be used for the  
CC detection of polynucleotides encoding human tubby homologue. The immune  
CC disorders that can be treated include AIDS, Addison's disease,  
CC adult respiratory distress syndrome, allergies, anaemia, asthma,  
CC atherosclerosis, Crohn's disease, ulcerative colitis, atopic dermatitis,  
CC gout, Grave's disease, irritable bowel syndrome, lupus erythematosus,  
CC multiple sclerosis, myasthenia gravis, osteoarthritis, osteoporosis,  
CC rheumatoid arthritis, scleroderma, and autoimmune thyroiditis. Also,  
CC complications of cancer, haemodialysis, extracorporeal circulation,  
CC infection, and trauma can be treated using the antagonist. The neuronal  
CC disorders that can be treated include Alzheimer's disease, amnesia,  
CC catatonias, amyotrophic lateral sclerosis, dementia, depression,  
CC Down's syndrome, epilepsy, Huntington's disease, multiple sclerosis,  
CC neurofibromatosis, Parkinson's disease, paranoid psychoses, schizophrenia  
CC and Tourette's syndrome.  
XX  
XX Sequence 372 AA;

Query Match 100.0%; Score 2020; DB 20; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.4e-204; Mismatches 0; Indels 0; Gaps 0;  
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVLSPLSFVKSHVRMOGILLIVFAKYQHLPYIQILSTKSTPTGLFGWNGKGVNICKL 60  
DB 1 mdvlsplsfkshvrmoqgilllvfakychlpqiqlstktstptglfgwngkgvnckl 60  
QY 61 KLYGYVSIINCHLPHISNNVORLEHFRILEMQNCEGRDTPNLDHLLIWFQDMNFR 120  
DB 61 klygyvsiinchlphissnnvorlehfrilemqncegrdtpnldhlliwfgdmnfr 120  
QY 121 IEDFGHVFRESIKNCYGLWEKDQSLAKKHDPQLLREFQGRLLFPPTYKFDNRNDY 180  
DB 121 iedfghlvfresikncycyglwekdqslakkhdpqlrllrefqgrllfpptykfdnrndy 180  
QY 181 DTSEKKRPKPAWTDRLILWLKRPCAGPDTPIPPASHFSLSLRGYSYSHMTYGISDHPKPSG 240  
DB 181 dtsekkrkpawtdrlilwlkrcpcaqdpdpippashfslslrgysyshmtygisdhpkpsg 240  
QY 241 TFDLELKPVSAPLIVLMPEDLWTVDNMVSYSTSDTFSPSPWDWIGLYKVGRLRVNDY 300

Db 241 tfdleikplvsapliivlmpedlwtvendnmvsysstsdtfspspwdwigykvglrvndy 300  
QY 301 VSYAWGDSKVSQSDNLNOVYIDISNIPPTTEDEFLCLCYNSLSRVVGISRPFIIPGSL 360  
DB 301 vsyawgdsksqsdnlnovyidisnipppttedeflclcynslsrvvgisrpfipgsl 360  
QY 361 REDPLGEAQFQI 372  
DB 361 redplgeaqpqi 372  
RESULT 2  
AAG73981  
ID AAG73981 standard; Protein; 279 AA.  
XX  
XX AAG73981;  
AC  
XX  
XX 03-SEP-2001 (first entry)  
DT  
XX  
XX Human colon cancer antigen protein SEQ ID NO:4745.  
DE  
XX  
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX W0200122920-A2.  
PN  
XX  
XX 05-APR-2001.  
PD  
XX  
XX 28-SEP-2000; 2000WO-US26524.  
PF  
XX  
XX 29-SEP-1999; 99US-0157137.  
PR  
XX  
XX 03-NOV-1999; 99US-0163280.  
PR  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX Ruben SM, Barash SC, Birse CE, Rosen CA;  
PI WPI: 2001-235357/24.  
XX  
XX N-PSDB; AAH33412.  
DR  
XX  
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers -  
PT  
XX  
XX Claim 11; Page 6545-6546; 9803pp; English.  
PS  
XX  
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P). Where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing P.  
CC Inactive proteins or to supplement the patients own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated ps,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAG77789 represent sequences used in the exemplification of the  
CC present invention.  
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
XX  
XX Sequence 279 AA;

Query Match 33.8%; Score 683.5; DB 22; Length 279;  
Best Local Similarity 52.8%; Pred. No. 1.7e-63;

Matches 130; Conservative 39; Mismatches 74; Indels 3; Gaps 3;

QY 1 MDVLSPLFKVSHVRMGILLVFAKYQHLYPIQILSTKSTPTGLFGYWGNGGVNCL 60  
Db 10 mdalqpfvnlvssvmqgvlillfakyyhlpflrdvqtdctrtglggywngkgvsvrl 69  
QY 61 KLYGYVSIINCHLPPHISNNYQRLHFDRILEQNCQEGRDIPNILDHLLIWFQDMNFR 120  
Db 70 aafghmlcfInchlpahmdkaegkrdnfqtllsqfpggagildhdvlfvfgdlnfr 129  
QY 121 IEDFGLHFVRESIKNRCYCGLWEKDQLSIAKKHDPPLREFQEGRLLEFPPTYKFDNRNSDY 180  
Db 130 iesydlhfvkfaiadsqhlqwekdqlnmakntwplkfgqegplnfaptkfdvgtoky 189  
QY 181 DTSEKKRKPATWDRILWRLKRPCAGPDTPIPPASH-FSLSLRGYSSHMYTGISDHRKVS 239  
Db 190 dtsakrkpawtdrllwkvk-apgggp-spsgrkshrlqvtqhsyrmeytvsdhkpxv 247  
QY 240 GTFDLE 245  
Db 248 aqfilq 253

RESULT 3  
AAB27797  
ID AAB27797 standard; Protein; 255 AA.

XX AAB27797;  
AC  
DT 29-JAN-2001 (first entry)  
XX Human secreted protein #4.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.  
XX  
OS Homo sapiens.

XX WO200055199-A1.  
XX 21-SEP-2000.  
XX 09-MAR-2000; 2000WO-US06014.  
XX 12-MAR-1999; 99US-0124095.  
XX 11-JUN-1999; 99US-0138598.  
XX 03-DEC-1999; 99US-0168665.  
XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;  
XX WPI: 2000-572359/53.  
XX N-PSDB; AAC59218.  
XX Isolated nucleic acid molecule encoding a human secreted protein is  
XX used in preventing, treating or ameliorating a medical condition -  
XX  
XX Claim 11; Page 372-373; 433pp; English.

XX Sequences AAB27794-B27840 represent the amino acid sequences of 47  
XX human secreted proteins encoded by the genes AAC59215-C59261. The genes  
XX and proteins are useful for preventing, ameliorating or treating medical  
XX conditions, e.g. by protein or gene therapy. The genes are isolated from  
XX a range of human tissues disclosed in the specification. The nucleic  
XX acids, proteins, antibodies and (ant)agonists are useful in the  
XX diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
XX ovarian cancer, and other cancers of the adrenal gland, bone, bone  
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune

CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,  
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
XX Sequence 255 AA;

Query Match 31.6%; Score 637.5; DB 21; Length 255;  
Best Local Similarity 52.6%; Pred. No. 1.le-58;  
Matches 121; Conservative 37; Mismatches 69; Indels 3; Gaps 3;

QY 17 MQGILLVFAKYQHLYPIQILSTKSTPTGLFGYWGNGGVNCLKLYGYVSIINCHLPP 76  
Db 1 mqgvilllfaakyyhlpflrdvqtdctrtglggywngkgvsvrlaafghmlcfInchlp 60  
QY 77 HISNNYQRLHFDRILEQNCQEGRDIPNILDHLLIWFQDMNFRIDFGLHFVRESIKNR 136  
Db 61 hmdkaegkrdnfqtllsqfpggagildhdvlfvfgdlnfrfiesydlhfvkfaiads 120  
QY 137 CYGGLWEKDQLSIAKKHDPPLREFQEGRLLEFPPTYKFDNRNSNDYDTSEKKRKPATWDRIL 196  
Db 121 qlhqlwekdqlnmakntwplkfgqegplnfaptkfdvgtoknydtsakrkpawtdrill 180  
QY 197 WRLKRPCAGPDTPIPPASH-FSLSLRGYSSHMYTGISDHRKVPVSGTFDLE 245  
Db 181 wkvk-apgggp-spsgrkshrlqvtqhsyrmeytvsdhkpxvxaqfilq 228

RESULT 4  
AAB64662  
ID AAB64662 standard; Protein; 508 AA.

XX AAB64662;  
AC  
DT 26-MAR-2002 (first entry)  
XX Drosophila melanogaster polypeptide SEQ ID NO 20778.  
DE Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
KW Drosophila melanogaster.

OS  
PN WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US09231.  
XX 23-MAR-2000; 2000US-191637P.  
XX 11-JUL-2000; 2000US-0614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI: 2001-656860/75.  
XX N-PSDB; ABL08765.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
XX Disclosure; SEQ ID NO 20778; 2lpp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 508 AA;

Query Match 25.1%; Score 507.5; DB 22; Length 508;  
Best Local Similarity 34.3%; Pred. No. 1.7e-44;  
Matches 111; Conservative 62; Mismatches 130; Indels 21; Gaps 7;  
QY 2 DVLSPLSFIRKSHVRMGIILLFAKYOHLPYIQLSTKPTGLFGYWGNGKGVNICKL 61  
Db 106 ellrnydyavkteqmgdlmsvrrqhvehlqdieaeftg9g9wgnkgavsvrft 165  
QY 62 LYGVYYSIINCHPLPHISNNYORLEHFDRIEMQNCGRDIPNILDHDLIWFQDMNFR 121  
Db 166 lygcglafvahlatahdumderiedykqlenhhvhkryreiydhdyfwfgdlnfrl 225  
QY 122 E--DFGLHFVRESKNC--VGGLEWQDL-SIAKKHDPPLRFEQGRLLFPPTPKFDRNS 177  
Db 226 qgsdsste-vrelvrdesqhealrqdlyqvreksqlafqvqlerlpafptfkfregt 284  
QY 178 NDYDTSEKKRPAWTDRIWR--LKRQPCAGPDTPIPPASHSLSLRGYSSHMTYGISD 234  
Db 285 seydl---krpawtdrimyavqplnrq-----pgmqisiecckshplytisd 331  
QY 235 HKPVSGTFDELEKPLVSAPLIVLMPEDLWTVENDMVMYSSTSDPSPWDWIGLYKVL 294  
Db 332 hkpvtsdftiklypnrvapgvvfpislwkgidentveyhkgaeftdegsndwigifpsey 391  
QY 295 RDVNDVYSYAWVGDSKVSCLN 318  
Db 392 asiadyvayeyvnaqespssdsn 415  
RESULT 5  
AAB27845  
ID AAB27845 standard; Protein; 381 AA.  
XX  
AC AAB27845;  
XX  
DT 29-JAN-2001 (first entry)  
XX  
Sequence homologous to protein fragment encoded by gene 4.

Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
neurological disease; infection; human; secreted protein.

OS Homo sapiens.  
XX  
PN WO200055199-A1.  
XX  
PD 21-SEP-2000.  
XX  
PF 09-MAR-2000; 2000WO-US06014.  
XX  
PR 12-MAR-1999; 99US-0124095.  
PR 11-JUN-1999; 99US-0138598.  
PR 03-DEC-1999; 99US-0168665.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.

XX  
PI Rosen CA, Ruben SM, Komatsoulis G;  
XX  
DR WPI; 2000-572359/53.  
XX

PT Isolated nucleic acid molecule encoding a human secreted protein is  
PT used in preventing, treating or ameliorating a medical condition -  
XX  
PS Disclosure; Page 491-493; 433pp; English.  
XX

CC The invention relates to the isolation of genes AAC59215-C59261 encoding  
CC the human secreted proteins AAB27794-B27840. This sequence represents a  
CC peptide fragment homologous to the protein encoded by the gene given  
CC in the descriptor line. The sequence is a search result from a BLASTX  
CC homology search. The genes and proteins are useful for preventing,  
CC ameliorating or treating medical conditions, e.g. by protein or gene  
CC therapy. The genes are isolated from a range of human tissues disclosed  
CC in the specification. The nucleic acids, proteins, antibodies and  
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:  
CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the  
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,  
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such  
CC as myocardial ischaemias; (d) wound healing; (e) neurological diseases  
CC as e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such  
CC as viral, bacterial, fungal and parasitic infections.

XX  
SQ Sequence 381 AA;

Query Match 24.5%; Score 494.5; DB 21; Length 381;  
Best Local Similarity 45.5%; Pred. No. 2.6e-43;  
Matches 105; Conservative 31; Mismatches 90; Indels 5; Gaps 4;  
QY 126 LHFVRESIKNRCYGGLEWQDLSTIAKKHDPPLRFEQGRLLFPPTPKFDRNSNDYDTSEK 185  
Db 1 lhfvtfaidsdqihlwekdqlmakntwplkgfegpInfaptfkfdvgtkkydtsak 60  
QY 186 KRKPATWTDRIWLKRPQPCAGPDTPIPPASH-FSLSLRGYSSHMTYGISDHPKPVSGTFDL 244  
Db 61 krkpawtdrillwkvk-apgggp-spsgkrkshrlqvtqhsyrshmevtcvshkpvaaqfll 118  
QY 245 ELKPLVSAPLIVLMPEDLWTVENDMVMYSSTSDPSPWDWIGLYKVLGVNDVYSYA 304  
Db 119 qfafrdmpilrvleavawppegavvymetvfarsswdwiglyrvgrhckdyavv 178  
QY 305 WVGDSKVSCLNQNQYVIDISNTPTEDEFLCYNSLSRVSGISRPOI 355  
Db 179 wakhedv--dgntyqvtfseeslpkghgdilgyvshhslilgitpefqi 227

RESULT 6  
AAB62704  
ID ABB62704 standard; Protein; 357 AA.  
XX  
AC ABB62704;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 14904.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEXE ) PE CORP NY.